Supplementary Information for

Native American gene flow into Polynesia predating Easter Island settlement

Alexander G. Ioannidis[1,2*†], Javier Blanco-Portillo[2*], Karla Sandoval[2], Erika Hagelberg[3], Juan Francisco Miquel-Poblete[4], J. Víctor Moreno-Mayar[5], Juan Esteban Rodriguez-Rodriguez[2], Consuelo D. Quinto-Cortés[2], Kathryn Auckland[6], Tom Parks[6], Kathryn Robson[7], Adrian V. S. Hill[6,8], María C. Avila-Arcos[9], Alexandra Sockell[10], Julian R. Homburger[10], Genevieve L. Wojcik[10], Kathleen C. Barnes[11], Luisa Herrera[12], Soledad Berríos[12], Mónica Acuña[12], Elena Llop[12], Celeste Eng[13], Scott Huntsman[13], Esteban G. Burchard[13], Christopher R. Gignoux[11], Lucia Cifuentes[12], Ricardo A. Verdugo[12,14], Mauricio Moraga[12,15], Alexander J. Mentzer[6,16], Carlos D. Bustamante[10,17], Andrés Moreno-Estrada[2†]

[1] Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA

[2] National Laboratory of Genomics for Biodiversity (LANGEBIO) - UGA, CINVESTAV, Irapuato, Guanajuato 36821, Mexico

[3] Department of Biology, University of Oslo, 1050 Blindern, N-0316 Oslo, Norway

[4] Departamento de Gastroenterología, Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile

[5] National Institute of Genomic Medicine (INMEGEN), Mexico City 14610, Mexico

[6] Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

[7] MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK.

[8] The Jenner Institute, Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7DQ, UK.

[9] International Laboratory for Human Genome Research (LIIGH), UNAM Juriquilla, Queretaro 76230, Mexico

[10] Center for Computational, Evolutionary and Human Genomics, Stanford University, Stanford, CA 94305, USA

[11] Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Denver, CO, USA

[12] Human Genetics Program, Institute of Biomedical Sciences, Faculty of Medicine, University of Chile, Santiago 8380453, Chile

[13] Program in Pharmaceutical Sciences and Pharmacogenomics, Department of Medicine, University of California San Francisco, San Francisco, CA, USA

[14] Basic-Applied Oncology Department, Faculty of Medicine, University of Chile, Santiago 8380453, Chile

[15] Department of Anthropology, Faculty of Social Sciences, University of Chile, Santiago 8380453, Chile

[16] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

[17] Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA

*These authors contributed equally to this work.
†To whom correspondence should be addressed.

Table of Contents

Supplementary Discussion

On generation times

An average factor of 30 years per generation is used to convert generation times to historical periods in years. The figure of 30 years per generation has been found by many studies to hold in pre-industrial populations[77,78] (for instance, Tremblay et al. give 32 years/generation from 19th century rural Quebec parish records[79], Matsumura et al. give 29 years for 19th century census data on Greenland Eskimos[80], Helgason gives 30 years from Icelander pedigree data from 1742 – present[81], and Fenner gives 29 years for hunter-gatherers[82]). This figure reflects the generation time averaged across all children and across both male and female parents. It reflects the average age of a woman across all of her lifetime offspring until menopause. It also reflects the average age of a man across all of his lifetime offspring, who he may continue to father very late in life, particularly where increasing age contributes to greater resources, status, and number of partners. Indeed, average generation times have decreased with modernization, as such cultural practices have changed. Fenner finds the modern developed-nations sex-averaged generation time to be two years lower than the hunter-gather average[82].

Iterative ADMXTURE motivation

The clusters found in an ADMIXTURE analysis do not necessarily correspond to ancestral populations; for example, a cluster specific to a sampled population may be found when that population is bottlenecked[56] or, equivalently, oversampled relative to other populations. Such sampling-scheme-specific clusters can be an interpretive nuisance, especially in cases where the heavily-sampled population is itself an admixture between multiple ancestral populations of interest. The Rapanui specific clusters seen in the unsupervised ADMIXTURE analyses of Supplementary Figure 1-2 subsume not only Polynesian ancestry on the island of Rapa Nui, but also some Native American ancestry. This is a problem that has been well analyzed before[56,57]. The reason that a single cluster is found in this case is that the Rapanui population is quite bottlenecked, having passed through both its ancestral founding bottleneck and a 19th century bottleneck of only a few hundred individuals due to the depredations of Peruvian slavers and the introduction of several diseases[30]. This means that the precolonial components on the island—both the prehistoric Native American component and the Polynesian component—are shared very uniformly across its modern inhabitants. Thus, when this population is overrepresented in the unsupervised clustering analysis (through the inclusion of nearly as many Rapa Nui samples, 166, as all of the remaining Pacific islands combined, 188), a Rapa Nui-specific cluster is found to be a good fit to the data, minimizing the model's residuals for these individuals. The fact that this apparent Rapa Nui cluster subsumes both the Polynesian and prehistoric Native American components may have contributed to the inability of Fehren-Schmitz et. al to identify the Native American component on this island, since those authors had only two admixed Polynesian individuals as references outside of their Rapanui references[11]. We address these issues with our Iterative ADMIXTURE approach and ancestry-specific PCA and MDS as described in Methods.

Supplementary References

77. Jobling, M., Hurles, M. & Tyler-Smith, C. *Human Evolutionary Genetics*. (Garland Science, 2013).
78. Stoneking, M. *An Introduction to Molecular Anthropology*. (John Wiley & Sons, 2016).
79. Tremblay, M. & Vézina, H. New estimates of intergenerational time intervals for the calculation of age and origins of mutations. *Am. J. Hum. Genet.* **66,** 651–658 (2000).
80. Matsumura, S. & Forster, P. Generation time and effective population size in Polar Eskimos. *Proceedings of the Royal Society B: Biological Sciences* **275,** 1501–1508 (2008).
81. Helgason, A., Hrafnkelsson, B., Gulcher, J. R., Ward, R. & Stefansson, K. A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: Evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *Am. J. Hum. Genet.* **72,** 1370–1388 (2003).
82. Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128,** 415–423 (2005).
83. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506,** 225–229 (2014).
84. Lindo, J. *et al.* The genetic prehistory of the Andean highlands 7000 years BP though European contact. *Science Advances* **4,** eaau4921 (2018).
85. Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349,** aab3884–aab3884 (2015).
86. Schroeder, H. *et al.* Origins and genetic legacies of the Caribbean Taino. *Proc. Natl. Acad. Sci. U.S.A.* **115,** 2341–2346 (2018).
87. Skoglund, P. *et al.* Genomic insights into the peopling of the Southwest Pacific. *Nature* **538,** 510–513 (2016).
88. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522,** 207–211 (2015).
89. D'Arcy, P. The Chinese Pacifics: A brief historical review. *The Journal of Pacific History* **49,** 396–420 (2014).
90. Nei, M. *Molecular Evolutionary Genetics*. (Columbia University Press, 1987).

Supplementary Tables

Population	# Individuals	Array	Population	# Individuals	Array
Atayal (Taiwan)	10	Axiom LAT-1	Paiwan (Taiwan)	12	Axiom LAT-1
Atiu (Cook Islands)	10	Axiom LAT-1	Palliser (Mataiva)	10	Axiom LAT-1
Aymara (Arica, Chile)**	16	Axiom LAT-1	Pehuenche‡	12	Axiom LAT-1
Aymara (Puno, Peru)	61	Illumina MEGA† and Axiom LAT-1	Raivavae	9	Axiom LAT-1
Chilote‡	30	Axiom LAT-1	Rapa Iti	16	Axiom LAT-1
Colombia*	96	Illumina 610-Quad	Rapa Nui (1994)	86	Axiom LAT-1
Ecuador*	20	Illumina 610-Quad	Rapa Nui (2013)	80	Axiom LAT-1
Huilliche‡	20	Axiom LAT-1	Rarotonga (Cook Islands)	3	Axiom LAT-1
Kaweskar‡	4	Axiom LAT-1	Rimatara	10	Axiom LAT-1
Magdalena de Cao	19	Illumina MEGA	South Marquesas	15	Axiom LAT-1
Mangareva (Gambier)	11	Axiom LAT-1	Tahiti	13	Axiom LAT-1
Mapuche**	32	Axiom LAT-1	Tubuai	18	Axiom LAT-1
Mauke (Cook Islands)	10	Axiom LAT-1	Vanuatu	16	Axiom LAT-1
Mixe†	28	Illumina MEGA	Yamana‡	14	Axiom LAT-1
Mixtec†	16	Illumina MEGA	Zapotec†	57	Illumina MEGA
North Marquesas	25	Axiom LAT-1	Zenu†	19	Illumina MEGA

*Bryc, K. *et al.* Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 8954–8961 (2010).
**Verdugo, R. A. Chilegenomico. Available at: www.chilegenomico.cl
†Wojcik, G. L., *et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature*, **570**(7762), 514–518 (2019).
‡la Fuente, de, C. *et al.* Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci. U.S.A.* **115,** 201715688–E4012 (2018).

Supplementary Table 1. Genotyping array and number of samples for each indigenous population included in this study

The 799 genotyped samples from indigenous populations that are included in this study along with the genotyping arrays used. Populations marked with footnotes have been previously published as indicated. For indicated analyses the samples from Rapa Nui were subdivided into a population with no European ancestry (6) and a population with high European ancestry (10).

Island	Sampling Date
Atiu (Cook Islands)	1986
Mangareva (Gambier)	1988 - 1990
Mauke (Cook Islands)	1986
North Marquesas	1988 - 1990
Palliser (Mataiva)	1988 - 1990
Raivavae	1988 - 1990
Rapa Iti	1988 - 1990
Rapa Nui	1994 & 2013
Rarotonga (Cook Islands)	1986
Rimatara	1988 - 1990
South Marquesas	1988 - 1990
Tahiti	1988 - 1990
Tubuai	1988 - 1990

Supplementary Table 2. Sampling date for each Polynesian population included in this study

Population	# Individuals	Array or Whole Genome Sequence
Americans of European Descent (CEU)*	60	Sequence
China (CHB)*	30	Sequence
Colombia (COL)*	30	Sequence
England†	77	Affymetrix GeneChip 500K
France†	89	Affymetrix GeneChip 500K
Ireland†	60	Affymetrix GeneChip 500K
Italy†	20	Affymetrix GeneChip 500K
Japan*	30	Sequence
Papua New Guinea‡	8	Sequence
Peru (PEL)*	37	Sequence
Spain†	106	Affymetrix GeneChip 500K
Spain (IBS)*	30	Sequence
United Kingdom (GBR)*	30	Sequence
Vietnam (KHV)*	30	Sequence
Yoruba (YRI)*	60	Sequence

*(1000 Genomes) 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 526.7571 (2015): 68-74.
†(POPRES) Nelson, Matthew R., et al. The Population Reference Sample, POPRES: a resource for population, disease, and pharmacological genetics research. *The American Journal of Human Genetics* 83.3 (2008): 347-358.
‡(HGDP) Bergström, Anders, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* 367.6484 (2020).

Supplementary Table 3. Number of individuals for each reference population used in this study.

Sample	Population
Saki1	Saki Tzul[22]
LaGalgada	La Galgada[22]
BigBar	BigBar[23]
PtaStaAna	PtaStaAna[23]
Ayayema	Ayayema[23]
Aconcagua	Aconcagua[23]
SpCave	SpCave[23]
Lovelock1	Lovelock[23]
Lovelock2	Lovelock[23]
Lovelock3	LovelockYng[23]
Lovelock4	Lovelock[23]
Sumidouro4	LagoaSta[23]
Sumidouro5	LagoaSta[23]
Sumidouro6	LagoaSta[23]
Sumidouro7	LagoaSta[23]
Sumidouro8	LagoaSta[23]
USR1	USR1[23]
Anzick1	Anzick1[83]
IL2	RioUncallane[84]
IL3	RioUncallane[84]
IL4	RioUncallane[84]
IL5	RioUncallane[84]
IL7	RioUncallane[84]
939	939[85]
Taino	Taino[86]
IPK12	AncKaweskar[19]
IPK13	AncKaweskar[19]
IPY10	AncYamana2[19]
IPY08	AncYamana2[19]

Supplementary Table 4. Source of ancient genomes from the Americas that were compared to Polynesians.

Island	African	European	Melanesian†	Native American	Polynesian
Atiu	0.00%	2.29%	0.35%	0.03%	97.33%
**Mangareva*	0.00%	10.27%	0.09%	1.00%	88.73%
Mauke	0.00%	0.44%	0.01%	0.07%	99.50%
**North Marquesas*	1.00%	9.71%	0.48%	4.40%	84.33%
**Palliser (Mataiva)*	0.00%	5.20%	0.35%	1.21%	93.20%
Raivavae	0.00%	0.00%	0.00%	0.00%	100.00%
Rapa Iti	0.00%	5.81%	0.76%	0.29%	93.06%
**Rapa Nui 1994*	0.06%	16.60%	0.14%	10.09%	73.10%
**Rapa Nui 2013*	0.26%	33.01%	0.21%	18.13%	48.35%
Rarotonga	1.60%	8.33%	0.40%	0.00%	89.67%
Rimatara	0.00%	4.00%	0.00%	0.01%	96.00%
**South Marquesas*	2.25%	13.20%	0.79%	3.86%	79.87%
Tahiti	0.00%	10.62%	3.19%	0.49%	85.77%
Tubuai	0.03%	13.72%	0.12%	0.01%	86.17%

Supplementary Table 5. Ancestry proportions from ADMIXTURE

Average ancestry proportions for each eastern Polynesian island in the dataset as determined by global ancestry clustering using an unsupervised (iterative) ADMIXTURE analysis with the optimal $K = 5$ (Supplementary Figure 3-4). The global comparison populations that were found to have higher than a 95% proportion for a given cluster are named along the top row. The Native American proportion (and European proportion) in the more recent Rapa Nui sampling (2013) is observed to be larger than in the one generation prior sampling (1994). Islands with at least 1% Native American ancestry are indicated (*) in italics. Note that the Melanesian (†) cluster is anchored by present-day samples from Vanuatu (Fig. 1). This is an early modern Oceanian component, as that is the period of our study, and is not equivalent to the ancient Papuan ancestry component discussed in studies of ancient near Oceania,[87] similarly the European component above is anchored by present-day samples from Europe and is not equivalent to the ancient Yamnaya component described in Haak et. al[88]. A similar caveat holds for the Polynesian component above, versus the ancient Austronesian component[87] that contributed to it.

Island	African	European	Native American	Polynesian
Atiu	0.00%	4.43%	0.31%	95.27%
**Mangareva*	0.05%	12.84%	2.13%	84.99%
Mauke	0.21%	0.83%	0.31%	98.65%
**North Marquesas*	0.78%	12.91%	4.38%	81.94%
**Palliser (Mataiva)*	0.27%	8.34%	1.82%	89.57%
Raivavae	0.00%	2.49%	0.71%	96.79%
Rapa Iti	0.01%	6.68%	0.38%	92.93%
**Rapa Nui 1994*	0.23%	20.40%	9.99%	69.36%
**Rapa Nui 2013*	0.60%	33.41%	17.15%	48.84%
Rarotonga	2.03%	9.83%	0.37%	87.80%
Rimatara	0.02%	8.68%	0.55%	90.73%
**South Marquesas*	2.12%	17.01%	3.11%	77.78%
**Tahiti*	0.20%	12.80%	1.27%	85.73%
Tubuai	0.34%	10.88%	0.64%	88.13%

Supplementary Table 6. Ancestry proportions from RFMix

Average ancestry proportions for each eastern Polynesian island in the dataset as determined by local ancestry inference using RFMix with $K = 4$. Islands with at least 1% Native American ancestry are indicated (*) and in italics.

	South Nat. American	European	Polynesian	Central Nat. American
South Nat. American	0	16.4734	22.506	16.2738
European	16.4734	0	22.8476	16.293
Polynesian	22.506	22.8476	0	4.8474
Central Nat. American	16.2738	16.2933	4.8474	0

Supplementary Table 7. Variance of ancestry proportion log-ratios in Rapanui individuals. Variances of log-ratios of each possible pair of ancestry proportions across Rapanui individuals give an indication of whether those ancestries are inherited independently[27]. Individual ancestry proportions were determined by an ADMIXTURE analysis (Figure 1b). The ratio of central Native American ancestry fraction to Polynesian ancestry fraction shows the smallest variance, indicating that the central Native American ancestry is associated with the Polynesian ancestry (see also Supplementary Figure 12), and not the European. This suggests that the central Native American ancestry introgressed first into the Polynesian ancestry, independent of later European colonial admixture into Polynesia. The next smallest variance is of the south Native American ancestry fraction with the European ancestry fraction, reflecting the association between those two ancestries on Rapa Nui stemming from immigration of admixed Chileans of European and south Native American ancestry to the island. Because Native American – European admixture in the Americas (in this case in Chile) occurred much later than central Native American – Polynesian admixture, the variance is higher for the log-ratio of the European – south Native American fraction than for the log-ratio of the Polynesian – central Native American fraction.

	South Nat. American	European	Polynesian	Central Nat. American
South Nat. American	1	0.01727	0.008703	0.0177
European	0.01727	1	0.008397	0.01766
Polynesian	0.008703	0.008397	1	0.1106
Central Nat. American	0.0177	0.01766	0.1106	1

Supplementary Table 8. Compositional correlations of ancestry proportion log-ratios in Rapanui individuals.

The compositional correlations (see Online Methods) for each possible pair of ancestry proportions in the 166 Rapanui individuals. Individual ancestry proportions were determined by an ADMIXTURE analysis (Figure 1b). See also Supplementary Figure 7 for interpretation.

	Polynesian	European	Central Nat. American
Polynesian	0	23.047	1.508
European	23.047	0	22.573
Central Nat. American	1.508	22.573	0

Supplementary Table 9. Variance of ancestry proportion log-ratios in Rapanui without southern Native American ancestry.

Variances of log-ratios of each possible pair of ancestry proportions in Rapanui individuals who lack southern Native American ancestry for comparison with Supplementary Data Table 7. Individual ancestry proportions were determined by an ADMIXTURE analysis (Figure 1b). The ratio of the central Native American ancestry fraction to the Polynesian ancestry fraction shows the smallest variance (see also Supplementary Data Table 10), indicating that the central Native American ancestry is associated with the Polynesian ancestry (see also Figure 3b), and not the European ancestry. This suggests that the central Native American ancestry introgressed first into the Polynesian ancestry, independent of later European colonial admixture into Polynesia.

	Polynesian	European	Central Nat. American
Polynesian	1	0.00822	0.293
European	0.00822	1	0.00864
Central Nat. American	0.293	0.00864	1

Supplementary Table 10. Compositional correlations of ancestry proportion log-ratios in Rapanui individuals without southern Native American ancestry

The compositional correlations (see Online Methods) for each possible pair of ancestry proportions in the 44 Rapanui individuals who lack southern Native American ancestry. Individual ancestry proportions were determined by an ADMIXTURE analysis (Figure 1b). See also Figure 3b and Supplementary Data Table 9. The association of central Native American ancestry fraction with Polynesian ancestry fraction, and not European ancestry fraction, suggests that the central Native American ancestry introgressed first into the Polynesian ancestry, independent of later European colonial admixture into Polynesia.

	Atiu	Mangareva	Mauke	North Marquesas	Palliser (Mataiva)	Raivavae	Rapa Iti	Rapa Nui	Rarotonga	Rimatara	South Marquesas	Tahiti	Tubuai
Atiu		0	0	0	0	0	0	0	0	0	0	0	0
Mangareva	0		0	0	0	0	0	0.0318	0	0	0	0	0
Mauke	0			0	0	0	0			0	0	0	0
North Marquesas	0	0	0		0.0958	0	0	0.00301	0	0	0.00833	0	0
Palliser (Mataiva)	0	0	0	0.0958		0	0	0	0	0	0	0.0154	0.0222
Raivavae	0	0	0	0	0		0	0	0	0	0	0	0
Rapa Iti	0	0	0	0	0	0		0	0	0	0	0	0
Rapa Nui	0	0.0318	0	0.00301	0	0	0		0	0	0	0	0
Rarotonga	0	0	0	0	0	0	0	0		0	0	0	0
Rimatara	0	0	0	0	0	0	0	0	0		0	0.0538	0.0667
South Marquesas	0	0	0	0.00833	0	0	0	0	0	0		0	0
Tahiti	0	0	0	0	0.0154	0	0	0	0	0.0538	0		0.0128
Tubuai	0	0	0	0	0.0222	0	0	0	0	0.0667	0	0.0128	

Supplementary Table 11. European ancestry-specific IBD network probabilities

The probability that two individuals, one selected at random from each island, share an IBD segment of at least 7 cM in their European genomic segments. The number of samples used from each island group is given in Supplementary Data Table 1.

	Atiu	Mangareva	Mauke	North Marquesas	Palliser (Mataiva)	Raivavae	Rapa Iti	Rapa Nui	Rarotonga	Rimatara	South Marquesas	Tahiti	Tubuai
Atiu		0	0	0	0	0	0	0	0	0	0	0	0
Mangareva	0		0	0	0.0182	0	0.0114	0.0515	0	0	0.0121	0.021	0
Mauke	0	0		0	0	0	0	0	0	0	0	0	0
North Marquesas	0	0	0		0.0542	0	0.00521	0.0502	0	0	0.0306	0.0192	0.00463
Palliser (Mataiva)	0	0.0182	0	0.0542		0	0.0125	0.0355	0	0	0.0533	0	0.0111
Raivavae	0	0	0	0	0		0	0	0	0	0	0	0
Rapa Iti	0	0.0114	0	0.00521	0.0125	0		0.0109	0	0	0	0	0
Rapa Nui	0	0.0515	0	0.0502	0.0355	0	0.0109		0	0	0.0414	0.0204	0.00335
Rarotonga	0	0	0	0	0	0	0	0		0	0	0	0
Rimatara	0	0	0	0	0	0	0	0	0		0	0	0
South Marquesas	0	0.0121	0	0.0306	0.0533	0	0	0.0414	0	0		0	0.00741
Tahiti	0	0.021	0	0.0192	0	0	0	0.0204	0	0	0		0.00855
Tubuai	0	0	0	0.00463	0.0111	0	0	0.00335	0	0	0.00741	0.00855	

Supplementary Table 12. Native American ancestry-specific IBD network probabilities

The probability that two individuals, one selected at random from each island, share an IBD segment of at least 7 cM in their Native American genomic segments. The number of samples from each island group is given in Supplementary Data Table 1.

Population Location	European Last	Polynesian Last	Nat. American Last
North Marquesas	**-174.9280927**	-386.4661933	-391.5077327
South Marquesas	**-181.187967**	-317.1913373	-320.3384519
Palliser (Mataiva)	**-144.2670348**	-190.9747836	-240.6134258
Mangareva	**-102.4136429**	-111.421556	-110.5398137
Rapa Nui (No European Ancestry)	-108.6628571	-	-
Rapa Nui (High European Ancestry)	-270.4414272	**-188.0245765**	-271.5941058

Supplementary Table 13. Log likelihoods for Tracts models shown in Figure 6.

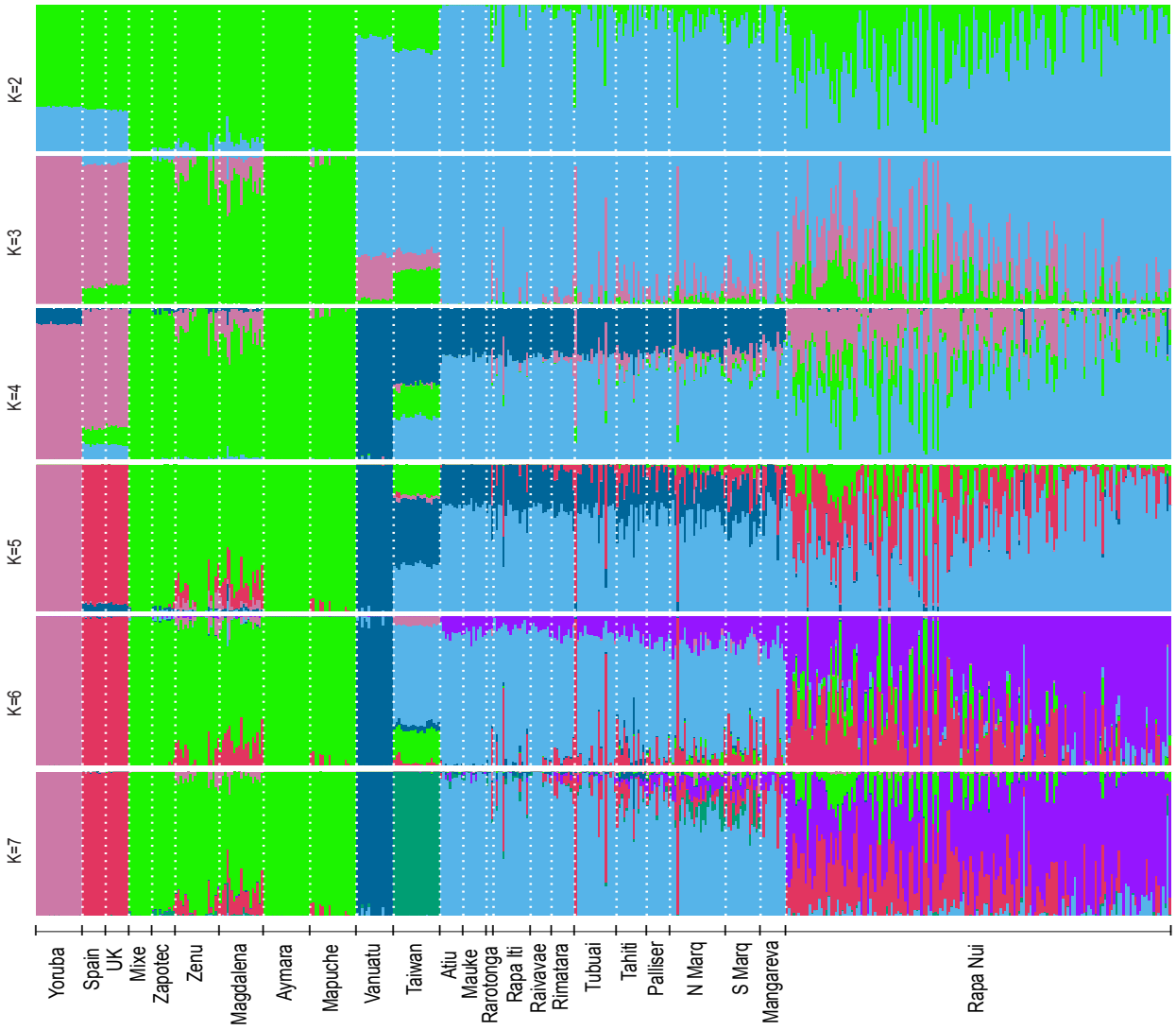
Log likelihoods for the tested Tracts models on each island (rows) with either the European, Polynesian, or Native American ancestries being the last to enter (cols). The best model likelihood for each island population is shown in bold, corresponding to the models plotted in Figure 6 a – f. The sample sizes for each island group are given in Supplementary Data Table 1.

# of Pulses	Proxy 1	Proxy 2	Amplitude 0	Time 0	Amplitude 1	Time 1	Amplitude 2	Time 2
1	Peru	Spain	7.24468e-05 +/- 4.65532e-06 (Z=15.5622)	15.4987 +/- 1.4535 (Z=10.6631)				
1	Peru	Taiwan	9.03666e-05 +/- 5.72291e-06 (Z=15.7903)	15.4987 +/- 1.4535 (Z=10.6631)				
2	Peru	Spain	7.96225e-05 +/- 4.09419e-06 (Z=19.4477)	30.0444 +/- 3.3539 (Z=8.95804)	1.57131e-05 +/- 4.66734e-06 (Z=3.3666)	5.07206 +/- 1.29407 (Z=3.91947)		
2	Peru	Taiwan	9.48255e-05 +/- 5.42983e-06 (Z=17.4638)	30.0444 +/- 3.3539 (Z=8.95804)	2.11366e-05 +/- 6.01124e-06 (Z=3.51618)	5.07206 +/- 1.29407 (Z=3.91947)		
3	Peru	Spain	7.80941e-05 +/- 4.10088e-06 (Z=19.0433)	30.4481 +/- 3.4483 (Z=8.82987)	0 +/- 2.66888e-07 (Z=0)	3.6073e-12 +/- 1.34245e-11 (Z=0.26871)	1.70521e-05 +/- 4.58452e-06 (Z=3.71951)	5.49158 +/- 1.36939 (Z=4.01024)
3	Peru	Taiwan	9.47923e-05 +/- 2.81828e-06 (Z=33.6349)	30.4481 +/- 3.4483 (Z=8.82987)	5.82869e-07 +/- 9.30538e-07 (Z=0.626379)	3.6073e-12 +/- 1.34245e-11 (Z=0.26871)	2.12506e-05 +/- 4.48523e-06 (Z=4.7379)	5.49158 +/- 1.36939 (Z=4.01024)

Supplementary Table 14. Multiple ALDER (MALDER) model.

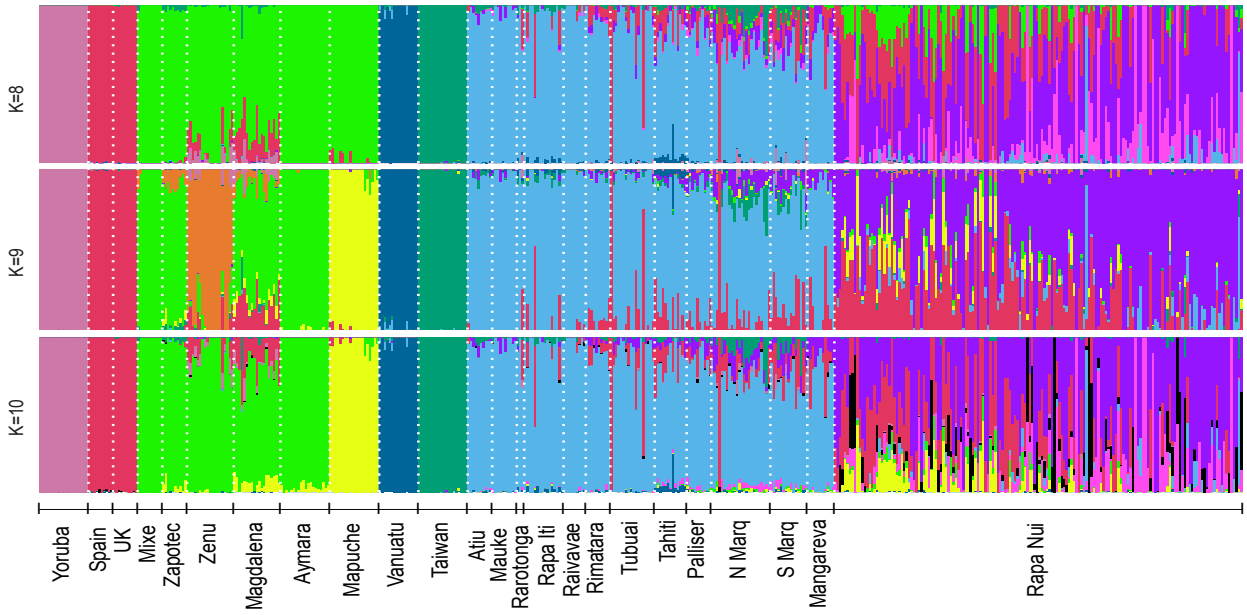
MALDER analysis of linkage disequilibrium decay in the 73 Rapanui (1994) individuals lacking African ancestry. Individuals (30) from Peru of only Native American descent, as indicated by ADMIXTURE, are used as references for the Native American component. Individuals (22) of indigenous Taiwanese descent (Atayal and Paiwan) are used as references for the unadmixed Austronesian (Polynesian) component. Individuals (30) of Spanish descent from 1000 Genomes are used as references for European descent. See Supplementary Data Tables 1-2. The discretized weighted LD was fit to exponential curves plus an affine term using least-squares to give the amplitudes and times, the standard errors were found by a jack-knife across the chromosomes and used for two-tailed Z-test values. An initial admixture event between Native Americans and Austronesians is found to have the highest amplitude. Several significant pulses of Austronesian-Native American admixture are found. The first pulse is dated to 30.4 +/- 3.4 generations, corresponding to 1082 CE (980 CE – 1184 CE) using a generation time of 30 years. A second pulse is dated to the present, corresponding to ongoing immigration of Native American descent individuals from modern Chile. A final pulse of 5.5 +/- 1.4 generations ago, or 1829 CE (1787 CE – 1871 CE), corresponds to the colonial period on Rapa Nui. When modeled with only two pulses the Native American introgression dates onto the island are 30 +/- 3.35 generations, or 1094 CE (994 CE – 1195 CE), and 5.1 +/- 1.3 generations, or 1841 CE (1802 CE – 1880 CE).

Supplementary Figures

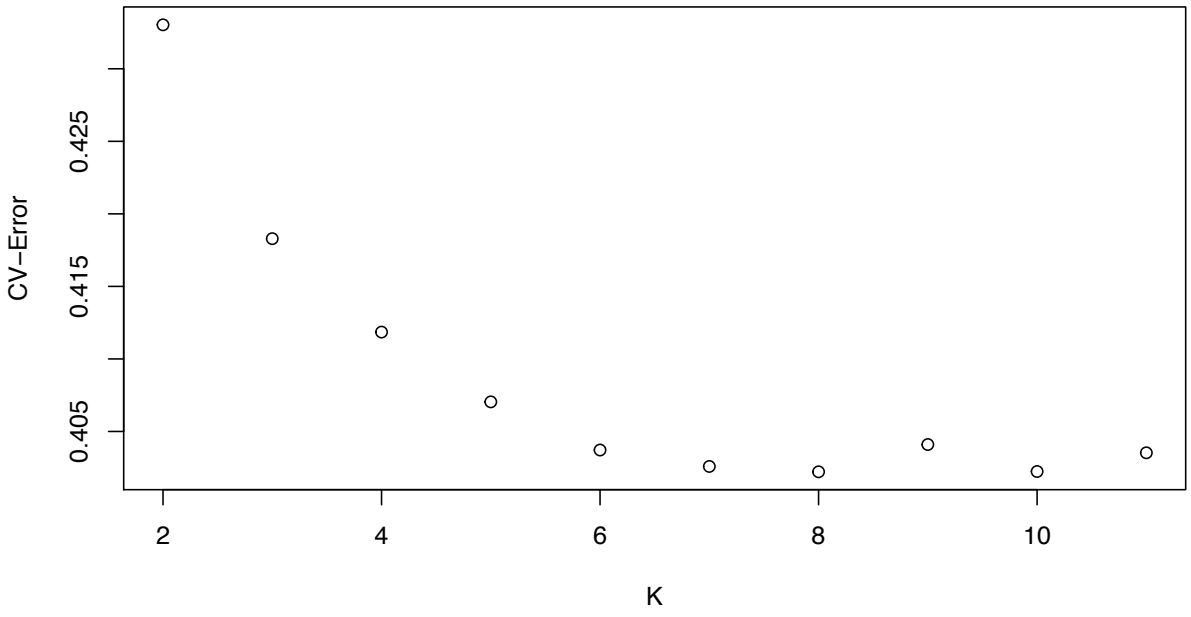


Supplementary Figure 1. Unsupervised ADMIXTURE results with K = 2-7 clusters.

ADMIXTURE was run on the intersection of Affymetrix (Axiom Lat-1) and Illumina (MEGA) SNPs (134,281 SNP overlap) from 489 samples. The populations are listed at bottom, and the numbers of samples from each are given in the Methods. Each column represents an individual with the length of each color bar representing the fraction of an ancestry cluster in the individual. The clustering is unsupervised, so the interpretation of each colored cluster must be inferred. Due to the large number of samples from the small island of Rapa Nui, a unique Rapanui component emerges at $K = 6$ (purple); the possibility that this component could subsume some of the islands' Native American ancestry, in addition to Polynesian, motivates our Iterative ADMIXTURE approach (Fig. 1). The other components at $K = 7$ are from left to right: African (mauve), European (red), Native American (green), Melanesian (dark blue), East Asian (olive), Polynesian (sky blue). Looking at $K = 7$ a Native American signature can be discerned in the five island populations at right: Rapa Nui, Mangareva, South Marquesas, North Marquesas, and Palliser. East Asian ancestry can be seen in the Marquesas, likely due to the documented 19th century immigration of Hakka laborers to that island group from south China[89].

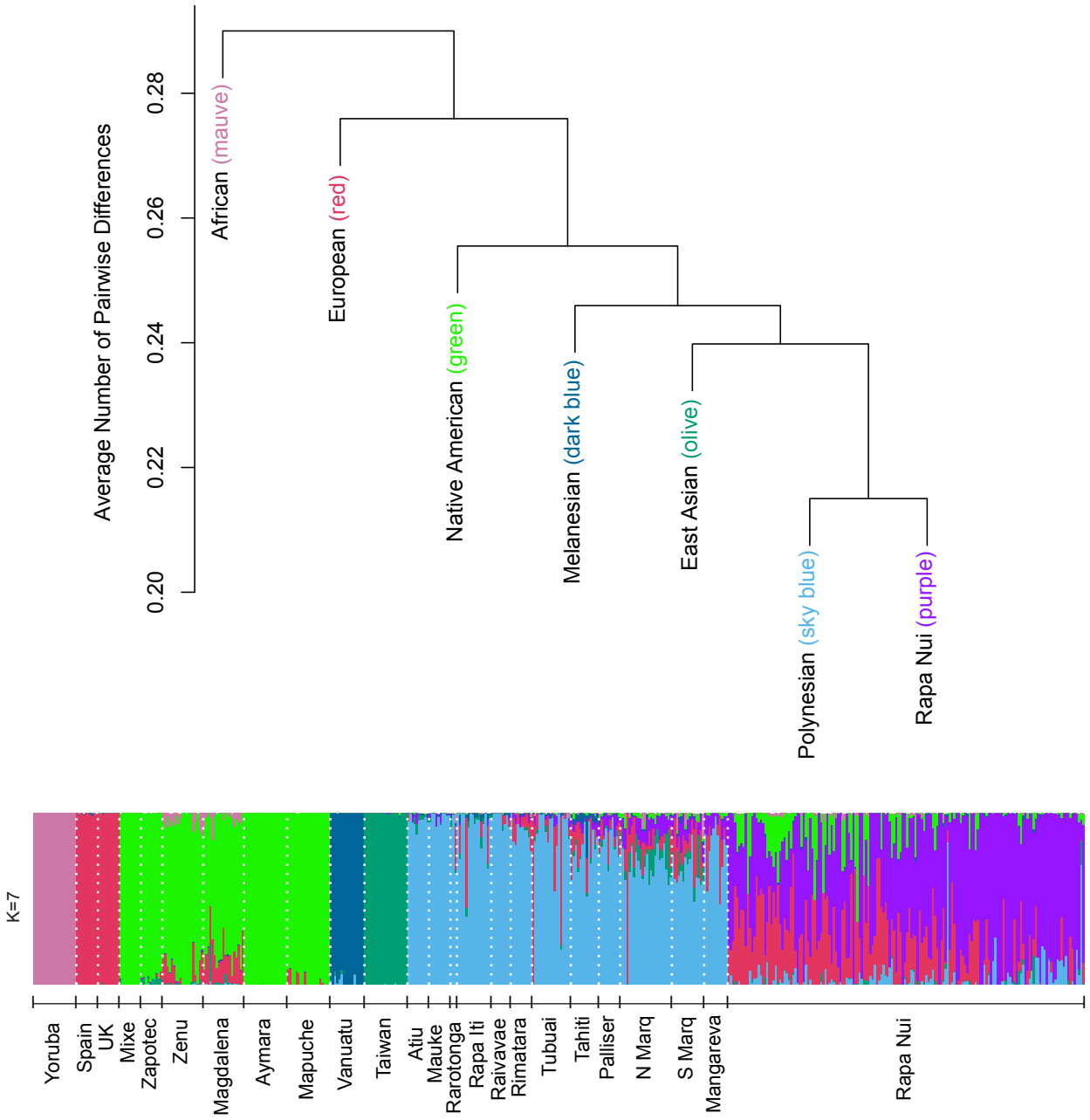


Supplementary Figure 2. Unsupervised ADMIXTURE results with K = 8-10 clusters. Additional unsupervised ADMIXTURE cluster plots (K = 8 – 10) for the 489 samples on the intersection of Affymetrix (Axiom Lat-1) and Illumina (MEGA) data (134,281 SNP overlap). The ADMIXTURE plots through the best fitting K = 7 (see Supplementary Figure 3) were plotted in Supplementary Figure 2. The populations are listed at bottom, and the numbers of samples from each listed population are given in Supplementary Data Tables 1-2. Each column represents an individual with the length of each color bar representing the fraction of an ancestry cluster in each individual. The clustering is unsupervised, so the interpretation of each colored cluster must be inferred. Due to the large number of samples from the population of Easter Island, the Rapanui, components characterizing the substructure of relatedness on this small island emerge at K = 8 (pink) and then again at K = 10 (pink and black). An earlier cluster (purple) also anchored by the overrepresented Rapanui emerged at K = 6 (Supplementary Figure 1). Such clusters are not particularly informative for determining the external origin of ancestries in the Rapanui and motivate our use of Iterative ADMIXTURE (Figure 1). Two Native American specific components also emerge. One (yellow) characterizes the southern Native American Mapuche population and another (orange) characterizes the Zenu group of Colombia.



Supplementary Figure 3. Cross-validation for ADMIXTURE analyses of Supplementary Figures 1-2.

An elbow[55.7] in the cross-validation error curve is seen at $K = 7$ clusters, indicating that this number of clusters provides a good model for fitting the data.

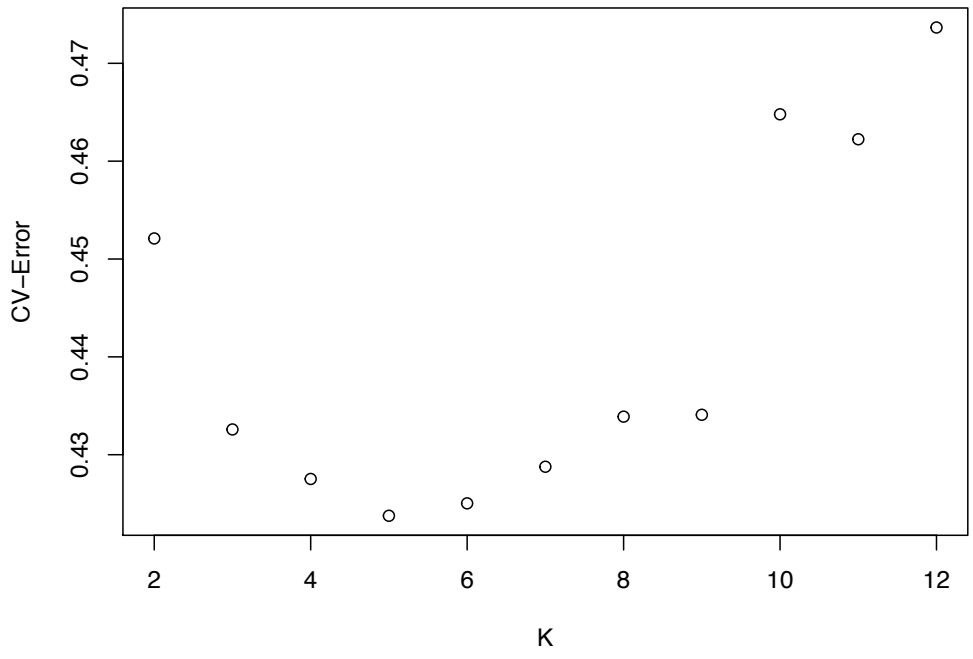


Supplementary Figure 4. Dendrogram for K = 7 ADMIXTURE analysis of Supplementary Figure 1.

Average linkage dendrogram constructed from average number of pairwise differences[90] between the seven clusters of the ADMXITURE analysis (K = 7) in Supplementary Figure 1 (reproduced above). The height of each split denotes the average number of pairwise differences between the joined clusters.

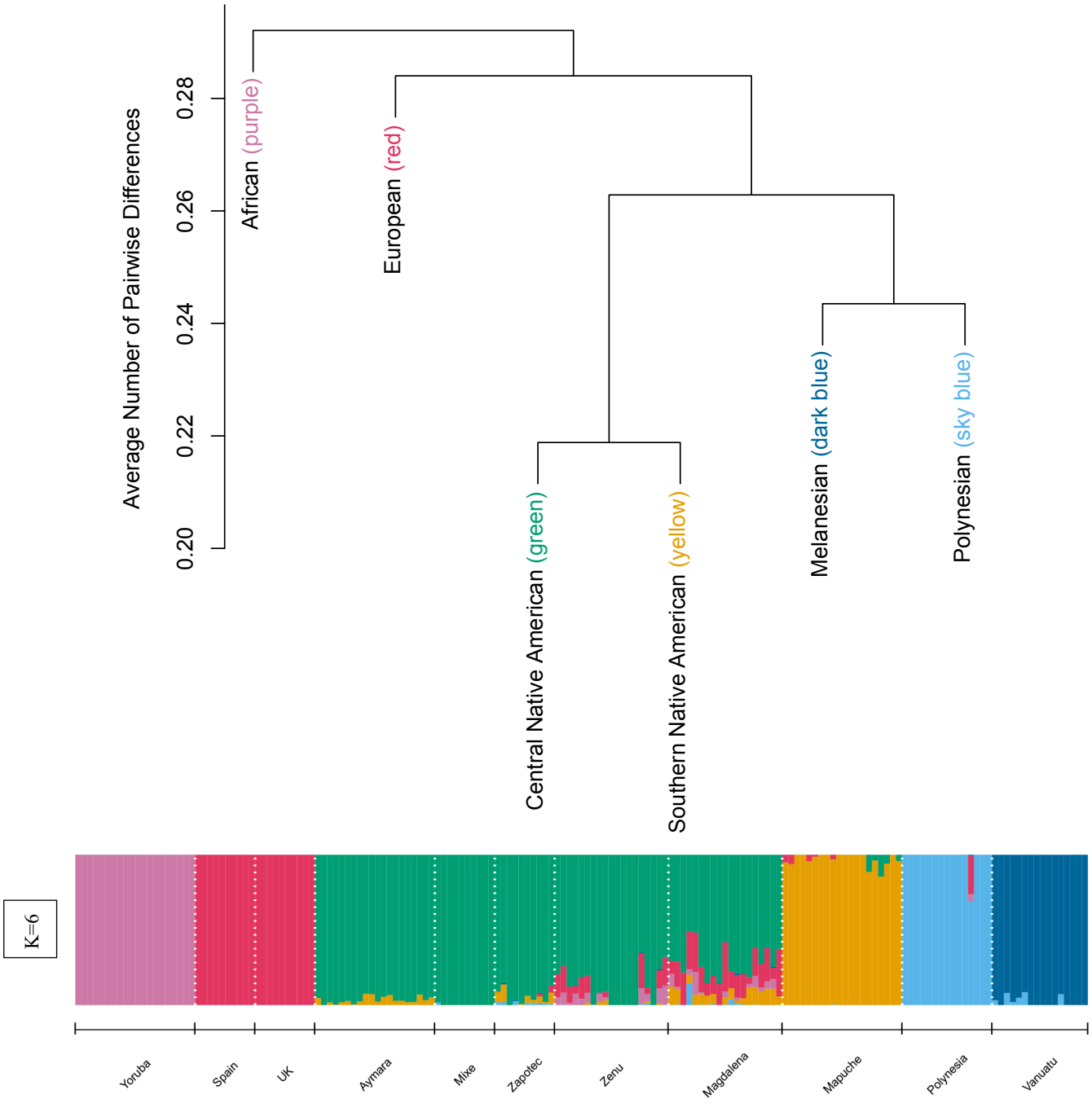
Supplementary Figure 5. Unsupervised ADMIXTURE results with K = 2-7 clusters used for Iterative ADMIXTURE

The populations are listed at bottom. There are 134,281 SNP markers, as in Supplementary Figure 1, and the numbers of samples from each listed populations are: 20 Yoruba, 10 Spain, 10 United Kingdom (UK), 20 Aymara, 10 Mixe, 10 Zapotec, 19 Zenu, 19 Magdalena, 20 Mapuche, 19 Polynesian, 16 Vanuatu. Each column represents an individual with the length of each color bar representing the fraction of an ancestry cluster in each individual. The clustering is unsupervised, so the interpretation of each colored cluster must be inferred. The clusters are from left: Africa (purple), Europe (red), central Native American (green), Zenu (orange), southern Native American (yellow), Polynesian (sky blue), and Melanesian (dark blue).



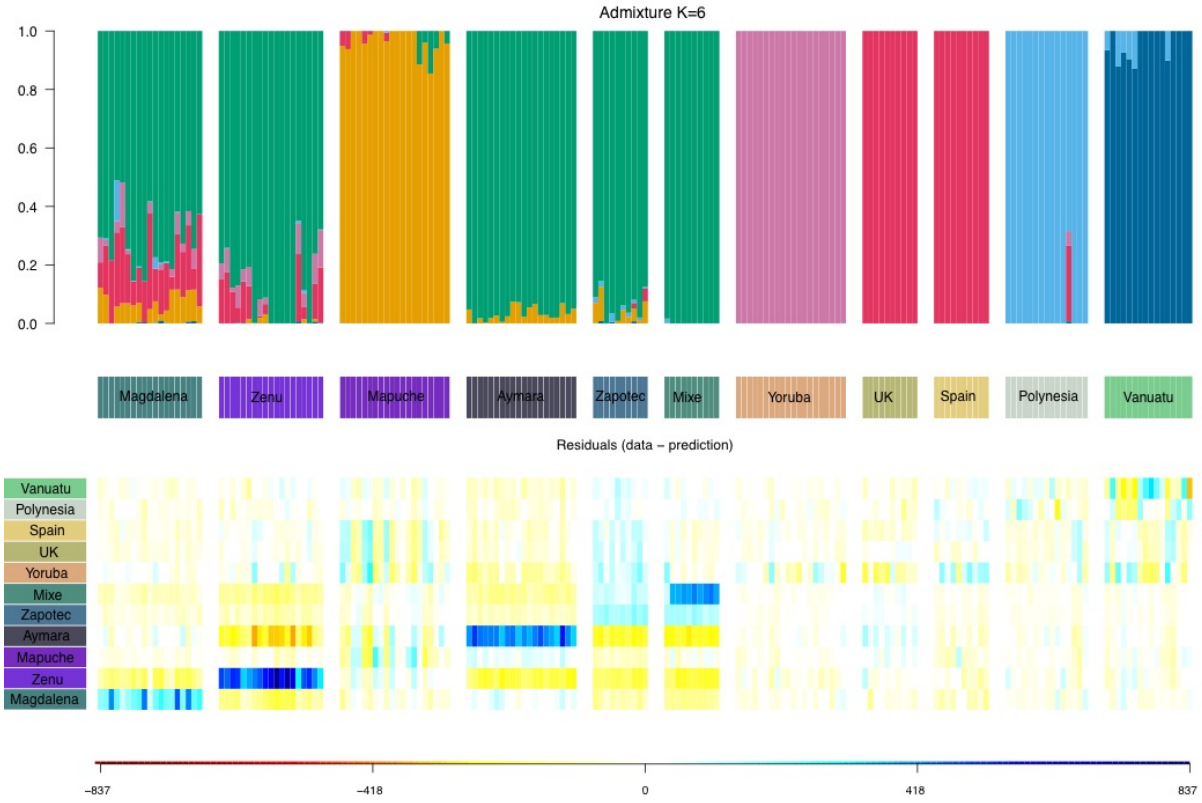
Supplementary Figure 6. Cross-validation for ADMIXTURE analysis in Supplementary Figure 5.

An elbow[55.7] in the cross-validation error curve is seen at $K = 5$ clusters. Using $K = 6$ clusters provides a similarly low cross-validation error and has the interpretative advantage of differentiating the southern Native American component (Chilean Mapuche) from the central Native American component (see Supplementary Figure 5).



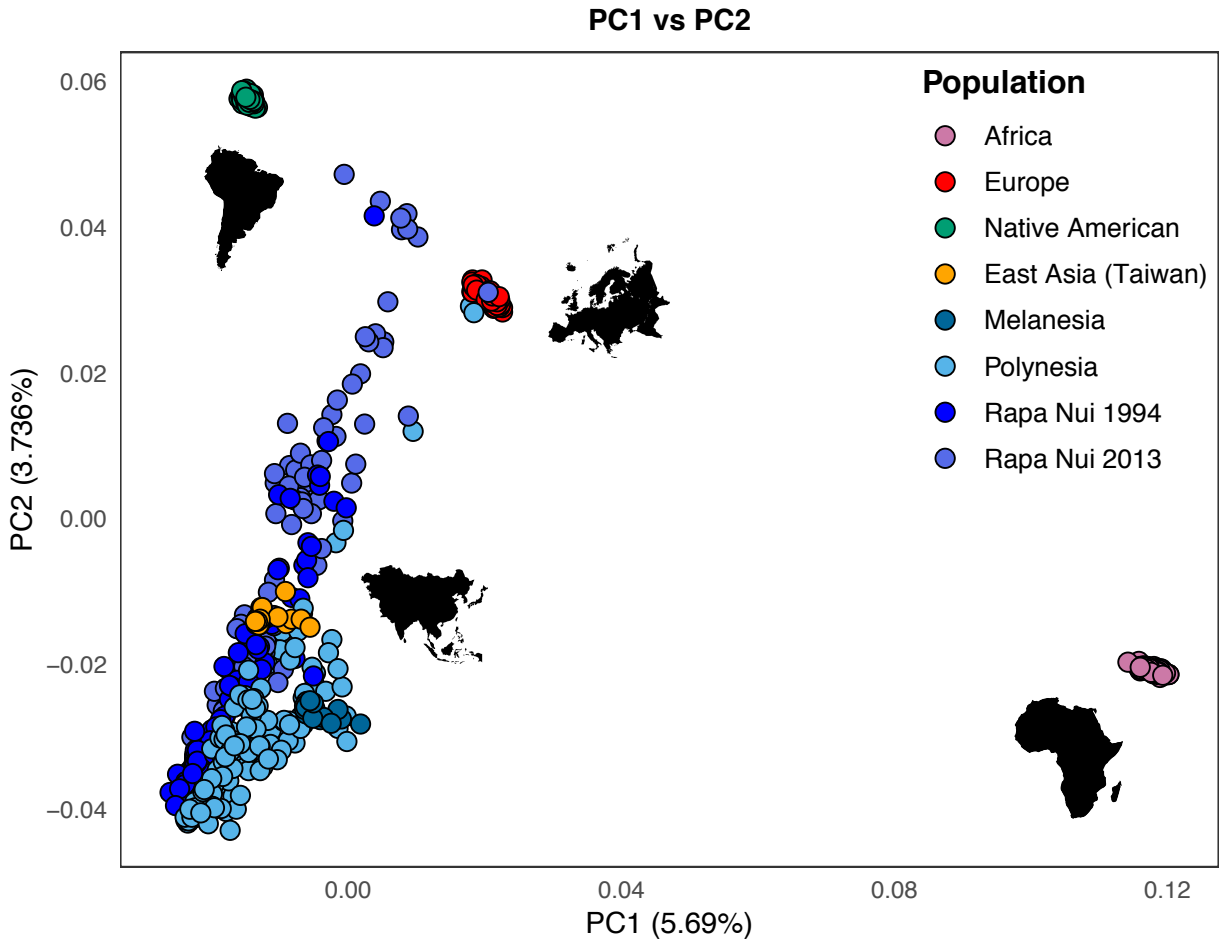
Supplementary Figure 7. Dendrogram for K = 6 ADMIXTURE analysis of Supplementary Figure 5.

Average linkage dendrogram constructed from average number of pairwise differences[90] between the six clusters of the ADMXITURE analysis ($K = 6$) in Supplementary Figure 5 (reproduced above). Height of each split denotes the average number of pairwise differences between the joined clusters.

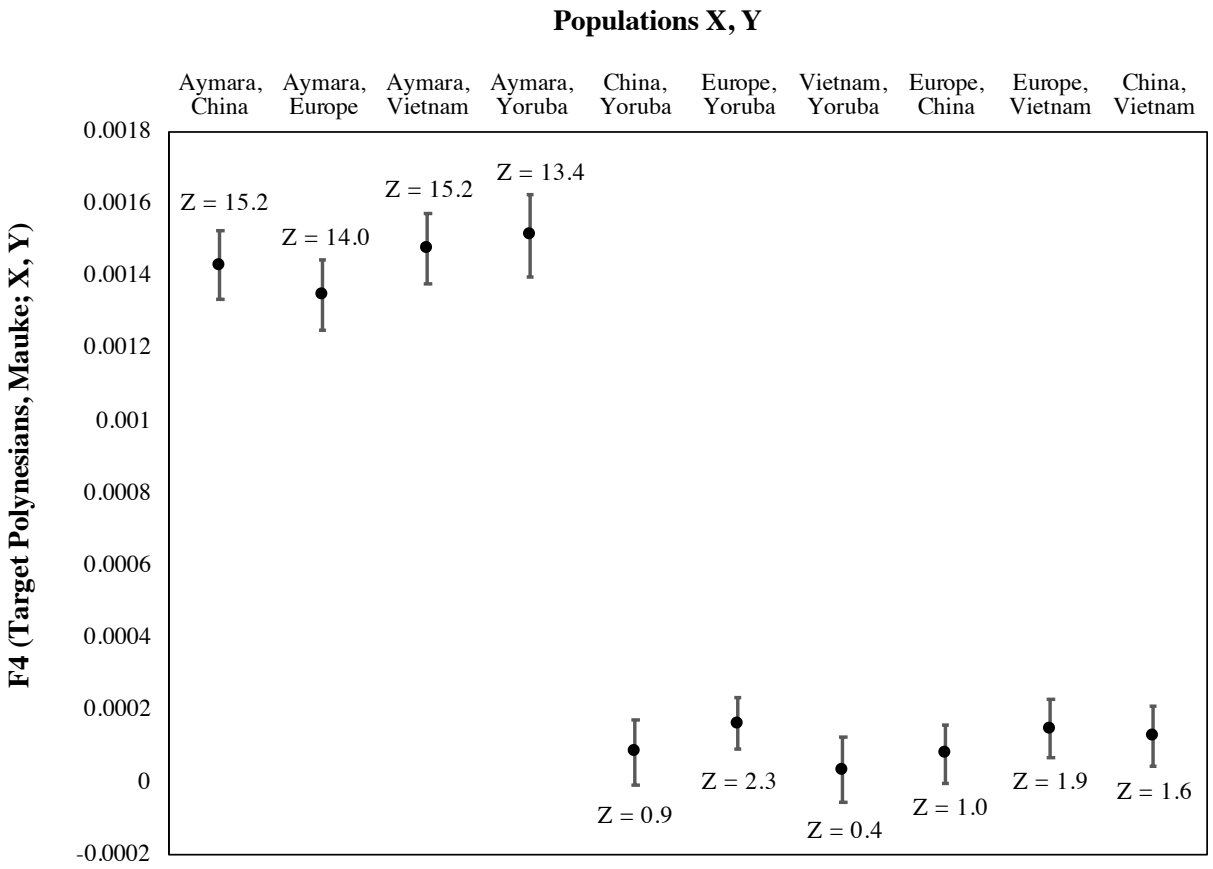


Supplementary Figure 8. Chromopainter based Badmixture analysis for K = 6 ADMIXTURE of Supplementary Figure 5 and 7.

A comparison of haplotype-sharing clusters between samples found using Chromopainter against the marker-frequency clusters found using ADMIXTURE via the method Badmixture [56]. At top is the *K* = 6 ADMIXTURE clustering of Figs 4 and 6 used as the basis for the Iterative ADMIXTURE analysis in Fig. 1. At bottom a heatmap grid shows how the empirical haplotype-sharing in these samples differs from their ADMIXTURE clustering; cool shades indicate more haplotype-sharing than expected and warm indicate less. Cool shades (more haplotype-sharing) are found along the diagonal for the Native American groups that were clustered in common (the green, central Native American ancestry cluster). This is because these groups are each isolated indigenous populations with recent haplotype sharing, due to relatedness, above that expected when considering them all a single ancestry cluster. The Zenu, a particularly isolated population, show strong, positive internal haplotype sharing deviations and strong, negative deviations from other co-clustered (green) Native American groups, particularly the Aymara. This presages the formation by the Zenu of a private ADMIXTURE cluster at the next higher *K* (cf. *K* = 7 Supplementary Figure 4). Importantly, off diagonal residuals between groups from different continental ancestry clusters (eg. European samples and Native American samples) are not prominent, indicating that those clusters are (from a haplotype sharing perspective) well separated.

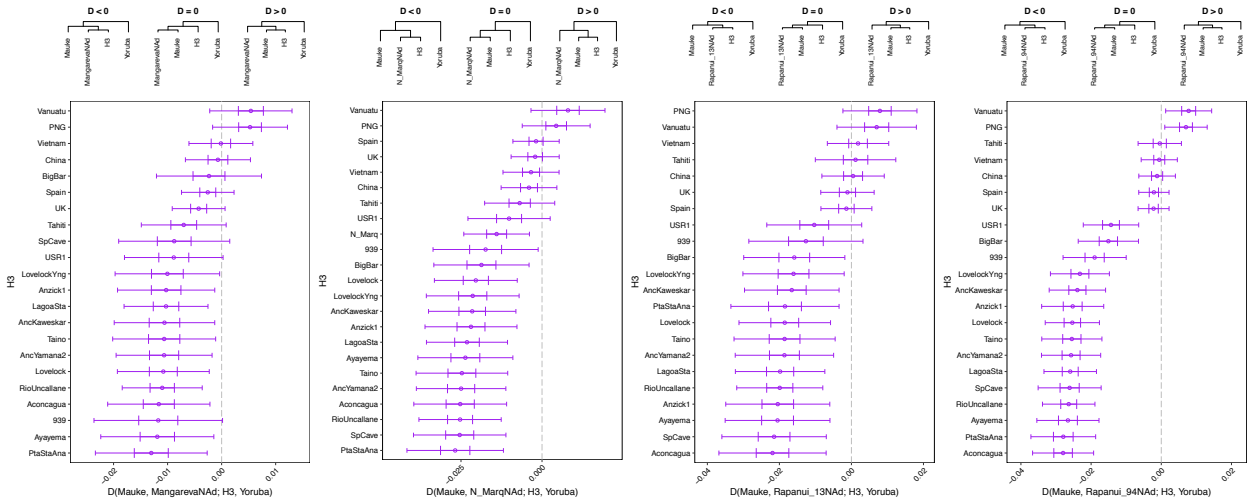


Supplementary Figure 9. Principal component analysis (PCA) of Pacific island samples.
All Pacific island samples are plotted in a principal component analysis of 689,899 SNP markers together with continental reference populations (labeled with maps, see also Supplementary Data Tables 1,3): Africa (60 Yoruba), Europe (30 UK, 30 Spain), East Asia (20 Taiwan), Native American (60 Puno), Melanesia (16 Vanuatu), Polynesia (150, all islands in Supplementary Data Table 2, excepting Vanuatu and Rapa Nui) and Rapa Nui (166, both 1994 and 2013). The Pacific island samples stretch toward a point between the Native American references and the European references, indicating that admixture from both of these ancestries is present in the Pacific islanders.



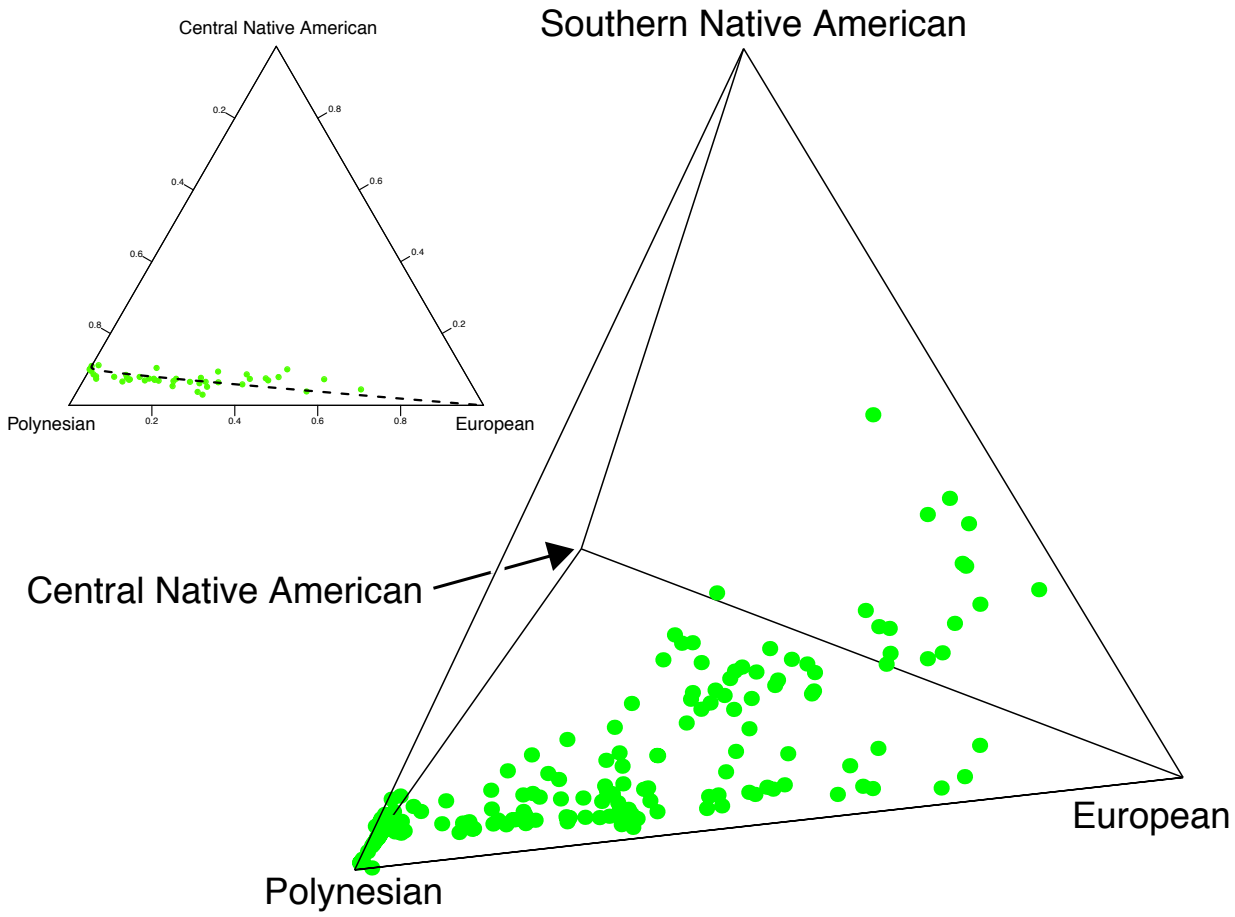
Supplementary Figure 10. F4-statistic test for admixture showing Native American introgression in Polynesia.

Target Polynesian individuals (the ten individuals having no European or African ancestry according to an ADMIXTURE analysis, Fig. 1) were chosen from the eastern islands previously found to have Native American introgression. To further test whether Native American ancestry exists in these samples, they were compared to the ten individuals from Mauke, a Polynesian island previously found to lack both European and Native American admixture. F4-statistics (points) and standard errors of them (bars) were computed between these two populations and various outgroups (X and Y), which are listed on the upper horizontal axis (see Supplementary Data Table 3 for reference sample sizes). If both outgroups contributed no admixture to either Polynesian population (Mauke or the target Polynesians), the F4-statistic will be zero, while if an outgroup contributed admixture to either of the two Polynesian populations, all F4-statistics involving that outgroup will be nonzero. F4-statistics involving Native Americans (the Aymara reference, top left) are all nonzero and significant, whereas F4-statistics involving only other outgroup population pairs (Europe, Africa, China, and Vietnam, top right) do not differ significantly from zero. (We followed the literature requiring $|Z|>3.3$, corresponding to a p-value<0.001, for rejecting a given null hypothesis[23]. This p-value for our two-tailed z-test is chosen to be conservative, since it does not include a multiple test correction.) This result confirms that neither the target eastern Polynesian individuals, nor the Mauke individuals, have European admixture (or East Asian admixture), and that there is Native American ancestry in either the Mauke or the target Polynesians. (Our ADMIXTURE and RFMix analyses confirm that this Native American admixture exists in the eastern Polynesian individuals, not Mauke.)



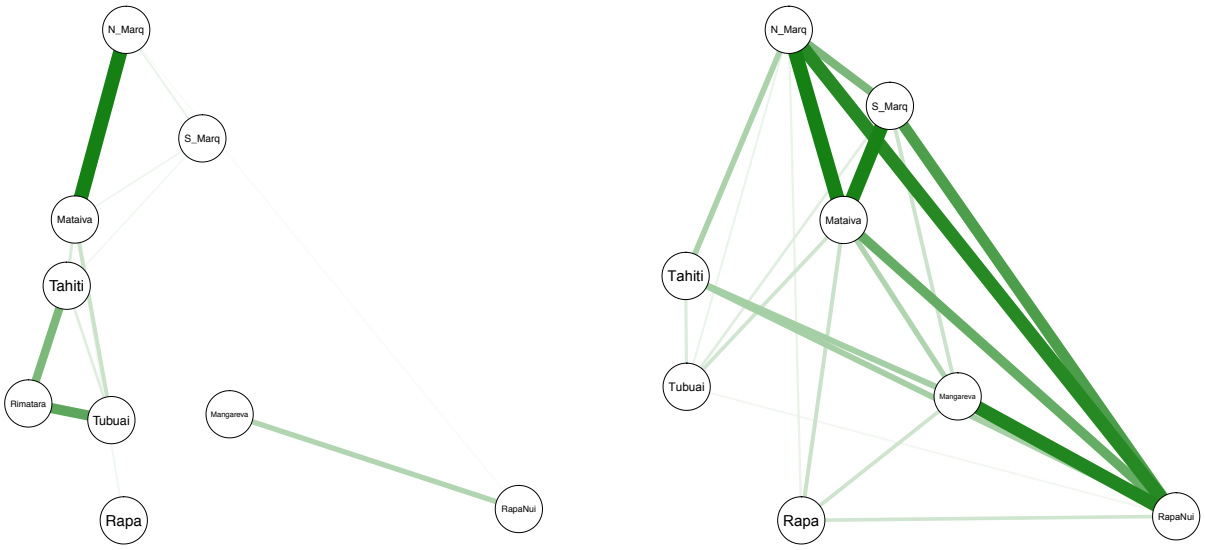
Supplementary Figure 11. D-statistic test for admixture, showing Native American introgression in Polynesia.

Target Polynesian individuals (having no European or African ancestry according to an ADMIXTURE analysis) were chosen from the eastern islands previously found to have Native American introgression: Mangareva (3), North Marquesas (1), Rapa Nui 2013 sampling (1), and Rapa Nui 1994 sampling (5). To further test whether these individuals carry Native American ancestry, we compared them to the ten individuals from Mauke, a Polynesian island previously found to lack both European and Native American admixture. We computed D-statistics of the form *D(Mauke, Target Polynesian; H_3, Yoruba)*. In this case, H_3 is one of a set of 22 tested reference populations and individuals (including the pre-contact ancient Native American genomes), as shown on the y-axis. Reference sample sizes are indicated in Supplementary Data Table 1,3, and 4. We expect that Native American admixture in the Target Polynesians will result in significantly negative values of D, when H_3 is Native American. A schematic representation of the null hypothesis (D=0) and the two possible outcomes of the alternative hypothesis (D<0 and D>0) for a two-tailed Z-test are shown at the top of each panel. Points represent estimated D-statistics, and nested error bars (estimated through a weighted block-jackknife procedure over 5-Mb blocks) represent one and 3.3 standard errors (corresponding to a p-value of 0.001 without multiple test correction). In agreement with previous results, we observe statistically significant deviations from $D=0$ ($|Z|>3.3$) when H_3 is Native American but not when H_3 is Eurasian. These results suggest that the test Polynesian individuals share significantly more alleles with Native Americans than Mauke islanders do. Moreover, such excess sharing appears to be independent of potential European-mediated admixture in modern Native Americans, since it holds when pre-contact ancient Native American individuals are used for H_3. The sharing is stronger with ancient Native American samples from South and Central America (and the Caribbean) than those from North America. The sharing is not related to the Papuan ancient component in Oceania, as the Papuans (PNG) and modern Vanuatu samples do not exhibit this signal of sharing with the target Polynesians.



Supplementary Figure 12. Compositional plot for ADMIXTURE ancestry proportions found in Rapa Nui.

Ancestry proportions from the ADMIXTURE analysis depicted in Figure 1 are plotted for the Rapa Nui samples. Each individual is represented by a point (green) at the location within the tetrahedron (lower right) corresponding to the ancestry fractions in that individual. Individuals with complete ancestry from any of the four sources would lie at the corresponding labeled vertex. Higher southern Native American ancestry individuals (top vertex) are seen to be associated with higher European component (right vertex). This is likely due to recent, differential introgression of an admixed (Spanish-southern Native American) Chilean component into islanders. Meanwhile, higher central Native American ancestry (back vertex) is associated with higher Polynesian ancestry (left vertex), rather than European; this can be more clearly seen in the ternary plot (Fig. 2b and reproduced in inset top left), which depicts only points lying on the base of the tetrahedron (samples with no recent southern Native American component). Samples in the inset have a clear association between their central Native American and Polynesian components (dotted line). In the tetrahedron, individuals lying above the base triangle rise off this association line in accordance with their southern Native American ancestry.

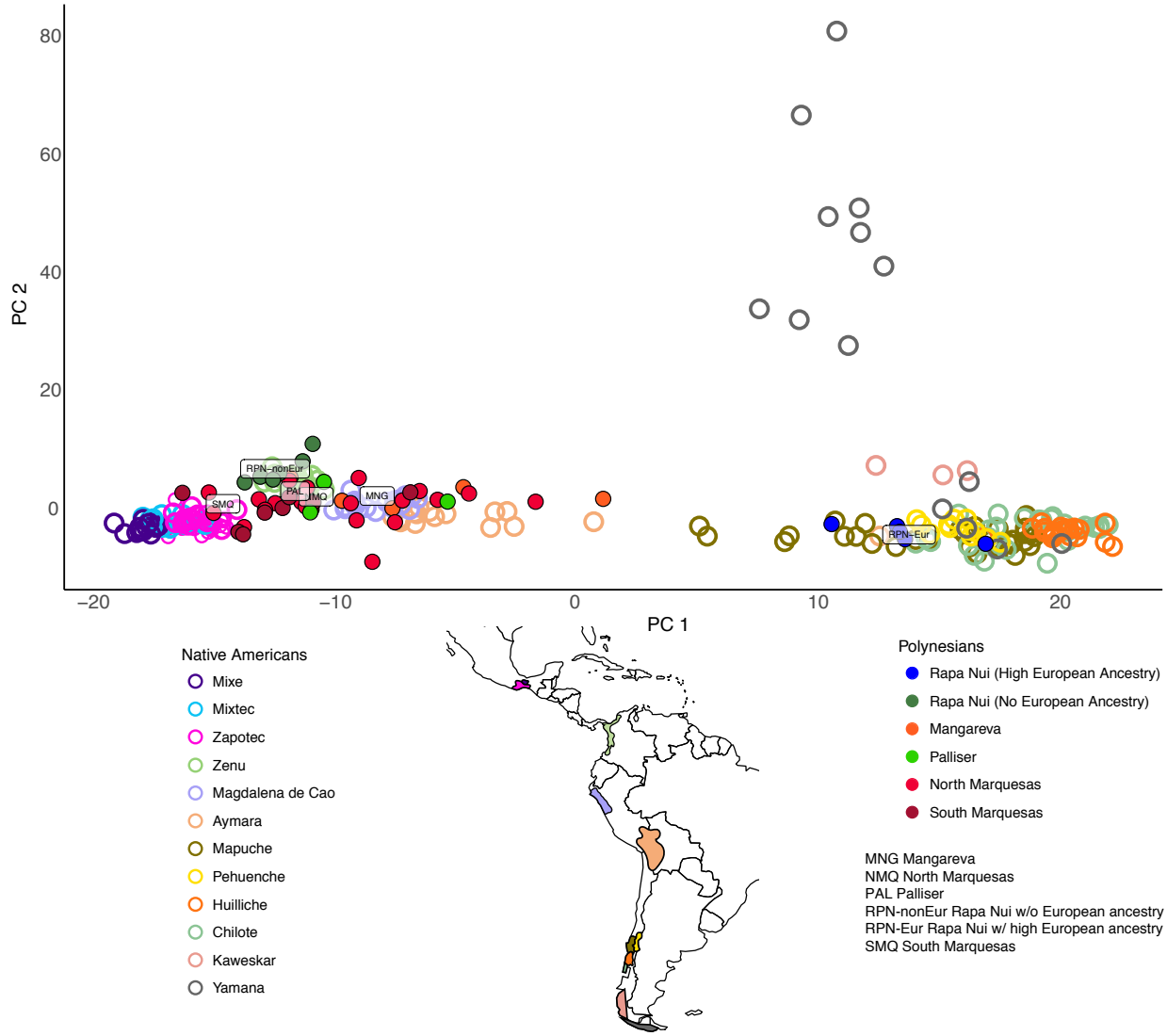


(a) European segments

(b) Native American segments

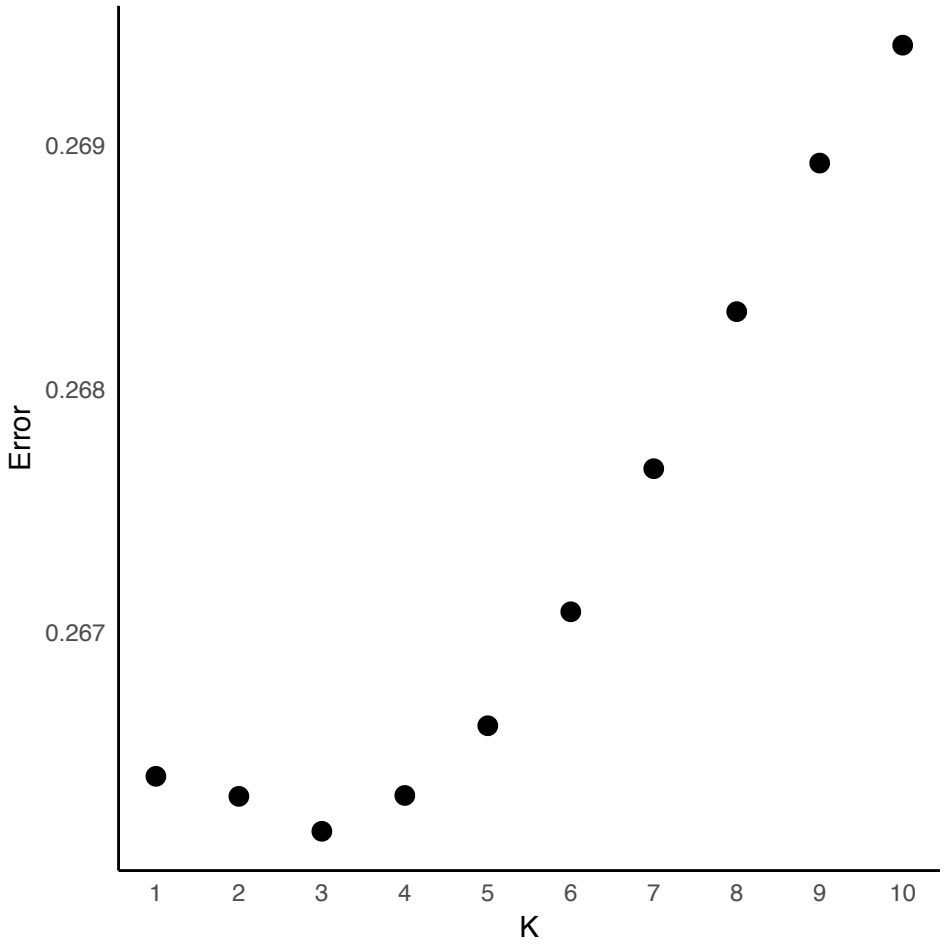
Supplementary Figure 13. Ancestry-specific identity-by-descent (IBD) networks in the European and Native American ancestry of Polynesians.

A side-by-side comparison of the IBD networks for segments found in the European (a) or Native American (b) genomic regions. Each connection has a width and shading proportional to the probability that two individuals, one selected at random from each island, share an IBD segment of over 7 cM in their European (a) or Native American (b) genomic segments. In the European segments (a) we see clustering along the secular boundaries of the European colonial process. The French speaking Polynesian islands (Tahiti, Rimatara, Tubuai, Rapa [Iti], Mataiva [Palliser], North and South Marquesas) form one connected component, while Rapa Nui, colonized by Spanish-speaking Chile, forms a separate component with a single connection to Mangareva. This connection could be explained by the evacuation from Rapa Nui in 1871 of the entire French Catholic mission led by Hippolyte Roussel to Mangareva (then possessing a population of 936 individuals) to escape the depredations of a French adventurer, who had seized Rapa Nui[28]. In the Native American segments (b) we see an entirely different pattern of inter-island IBD sharing, reflecting the different, pre-colonial, history of Native American contact in the eastern Polynesian islands. The sample sizes for each island are given in Supplementary Data Table 1.



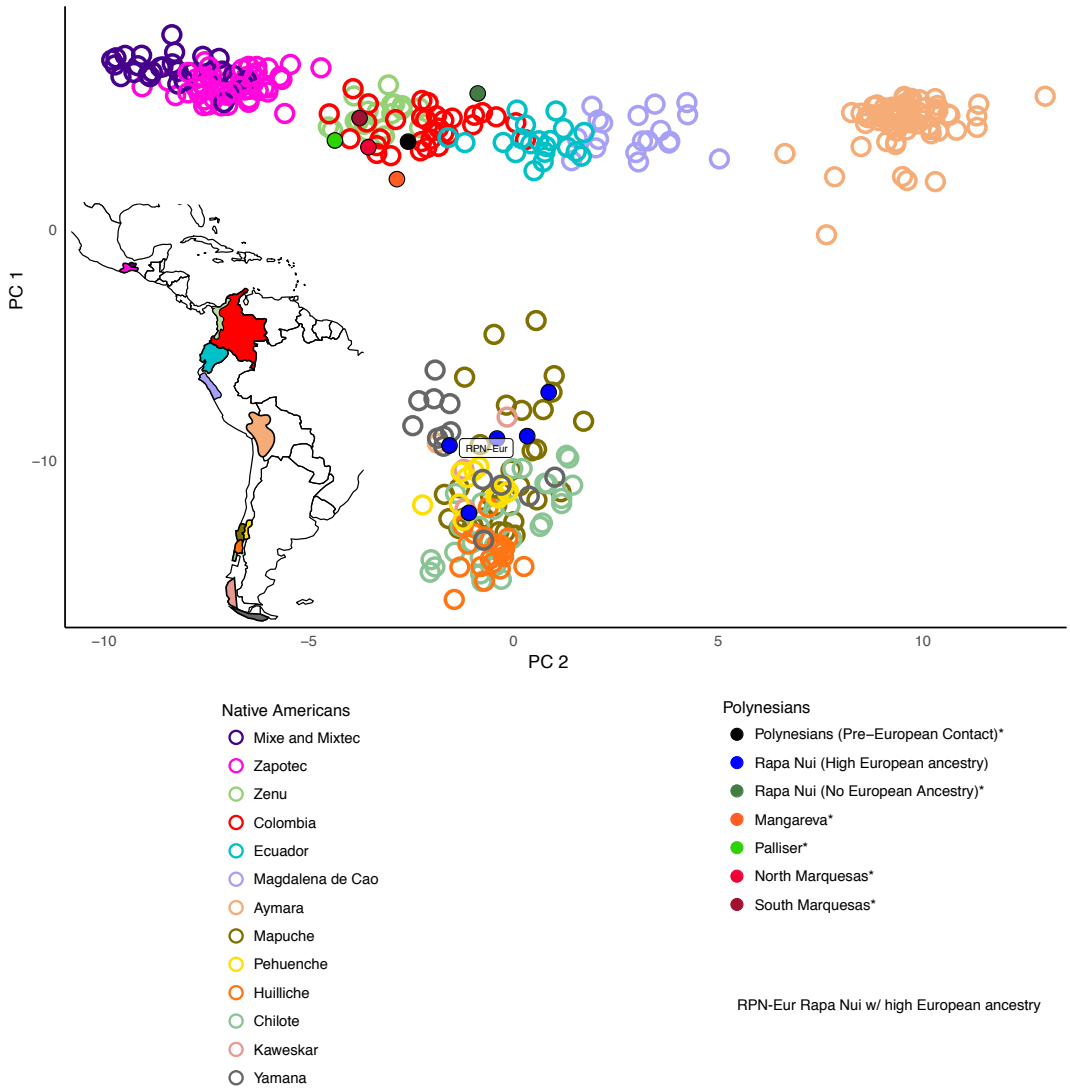
Supplementary Figure 14. Native American ancestry-specific PCA.

Our new weighted ancestry-specific SVD-completed principal component analysis (PCA) applied to the Native American component from Pacific islanders and reference individuals from the Americas. The Native American ancestries of reference individuals are plotted with open circles. The Native American ancestries of Pacific islanders are represented as solid points. Only individuals with at least 90,000 SNP markers in Native American tracts were plotted in this analysis. In addition, labels are plotted for the location of the allele frequency vector created from aggregating all Native American ancestry haplotypes in each island population. The Rapanui are split into two populations, those without European ancestry tracts (dark green) and those with high European and Native American ancestry (blue). The first principal component axis separates the reference Native Americans on a north-south axis, while the second principal component separates the Yamana of southern Patagonia. Pacific islanders' Native American ancestry is seen to cluster with northern South American references with the exception of the Rapanui with high European ancestry, who cluster with the southern references (Chilean). The population sample sizes for each indicated group are given in Supplementary Data Table 1.

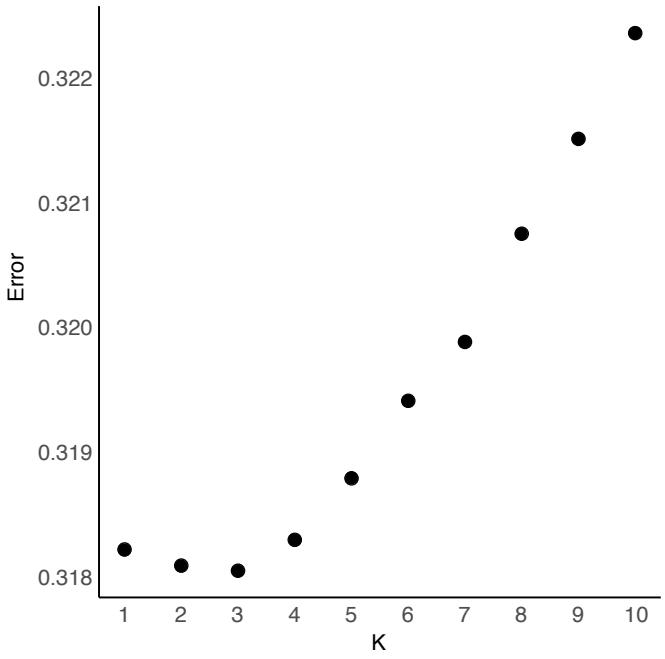


Supplementary Figure 15. Cross-validation plot for SVD completion of Native American ancestry-specific sample matrix.

The cross-validation plot for the singular value decomposition (SVD) completion of the Native American ancestry-specific masked sample matrix used to produce Supplementary Figure 14, showing a minimum reconstruction error at $K = 3$ dimensions.

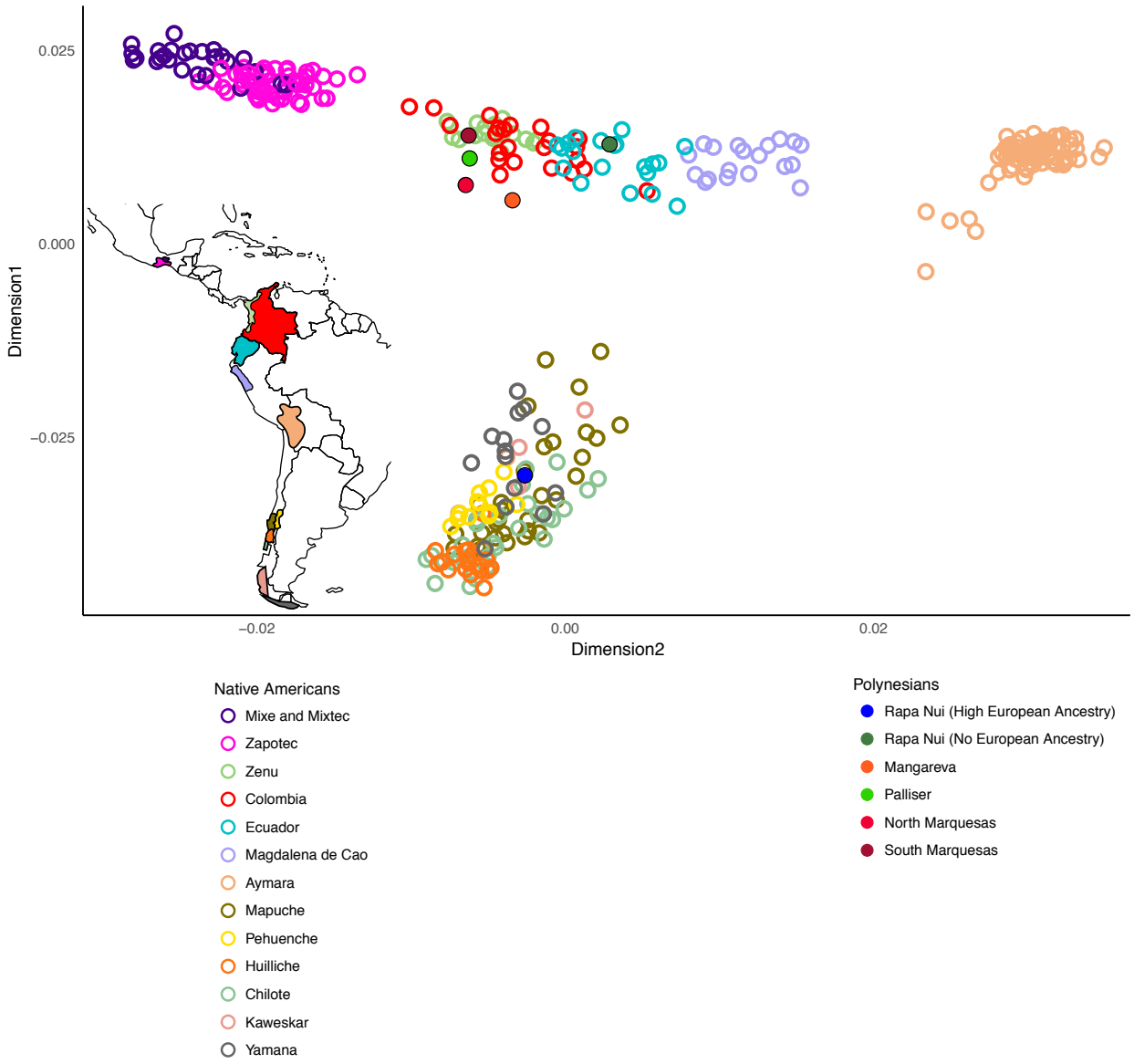


Supplementary Figure 16. Native American ancestry-specific PCA (additional references). Our new weighted ancestry-specific SVD-completed principal component analysis (PCA) applied to Native American ancestry from Pacific islanders and reference individuals from the Americas including additional Native American populations. The Native American ancestries of reference individuals are plotted with open circles. Insufficient markers remain in this intersection of three arrays to plot the Native American ancestry of most individual Pacific islanders. Instead, solid circles are plotted for the location of the allele frequency vector created from aggregating all Native American ancestry haplotypes in each island population (marked with * in legend). The Rapanui are split into two populations, those without European ancestry tracts (plotted as an aggregate point), and those with high European and Native American ancestry, who have sufficient Native American ancestry to be plotted individually as solid points (blue) and not as an aggregate frequency vector. Their aggregate frequency vector location is plotted as text. Pacific islanders' Native American ancestry is seen to cluster with references from Colombia, with the exception of the Rapanui with high European ancestry, who cluster with the references from Chile. An allele frequency vector for the Native American component combined across all islands (excepting high European ancestry Rapa Nui) is also shown (black). The population sample sizes for each indicated group are given in Supplementary Data Table 1.



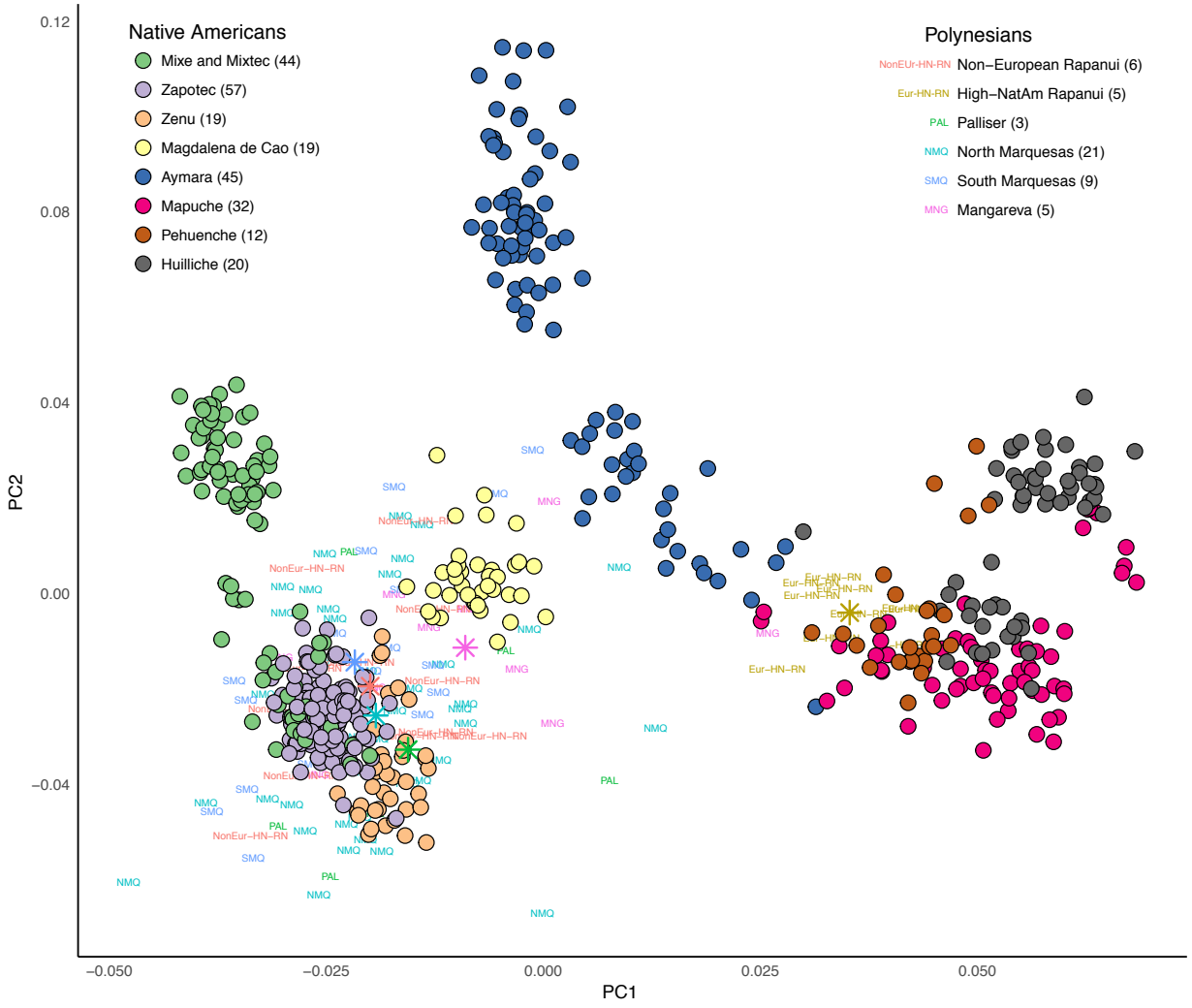
Supplementary Figure 17. Cross-validation plot for SVD completion of Native American ancestry-specific sample matrix with additional references.

The cross-validation plot for the singular value decomposition (SVD) completion of the Native American ancestry-specific masked sample matrix used to produce Supplementary Figure 16 showing a minimum reconstruction error at $K = 3$ dimensions.



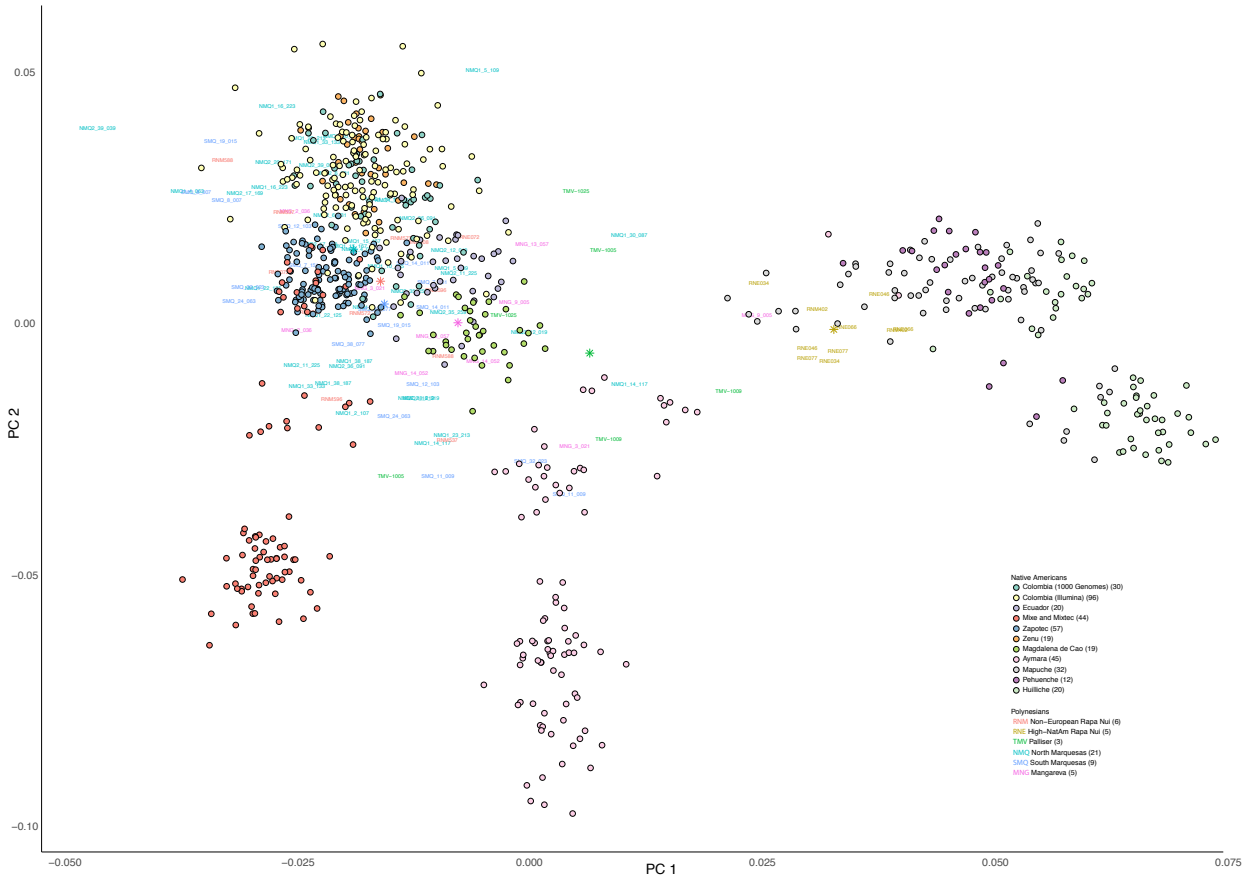
Supplementary Figure 18. Ancestry-specific MDS applied to the Native American component of Pacific islanders and reference individuals from the Americas.

Our new ancestry-specific MDS for comparison with the PCA in Supplementary Figure 16. The Native American ancestries of reference individuals (with post-colonial European and African admixture removed) are plotted with open circles. Solid circles are plotted for the location of the allele frequency vector created from aggregating all Native American ancestry haplotypes in each island population. The Rapanui are split into two populations, those without European ancestry tracts (dark green) and those with high European and Native American ancestry (blue). Pacific islanders' Native American ancestry is seen to cluster with references from Colombia with the exception of the Rapa Nui with high European ancestry, who cluster with the references from Chile. The sample size for each population is given in Supplementary Data Table 1.



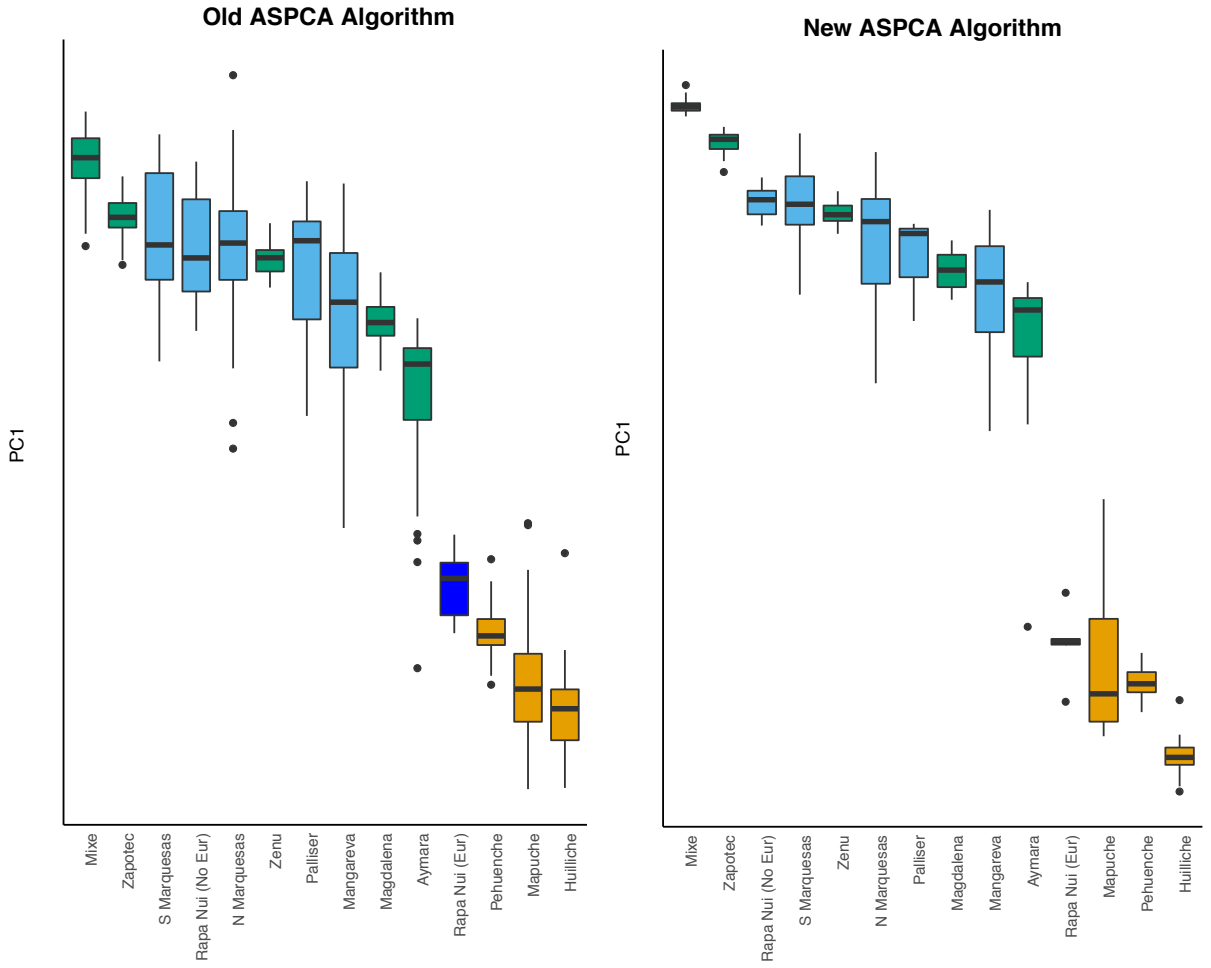
Supplementary Figure 19. ASPCA of the Native American component in Pacific islanders and references.

An ancestry specific PCA using the previously published[70] algorithm for samples genotyped on Illumina MEGA and Affymetrix Axiom-LAT 1 arrays (91,835 SNPs intersection). Each haploid genome is represented separately, since this previous method is haploid based. The Native American ancestry of reference individuals is plotted with colored points. The Native American ancestry of Pacific island individuals having greater than 2% Native American ancestry (as determined by local ancestry) is plotted with colored text. The centroid of the positions of the individuals from each island are plotted with stars matching the color of that island's text labels. With the exception of the Rapanui with high Native American ancestry (>40%), who cluster with the references from Chile, the Native American ancestry in Pacific island individuals is seen to cluster with northern South American and Central American references. For the low Native American ancestry individuals from the Pacific islands, the noise (scatter) due to the low number of SNPs available in the intersection is evident. The sample sizes for each group are given in parentheses after the group's name.



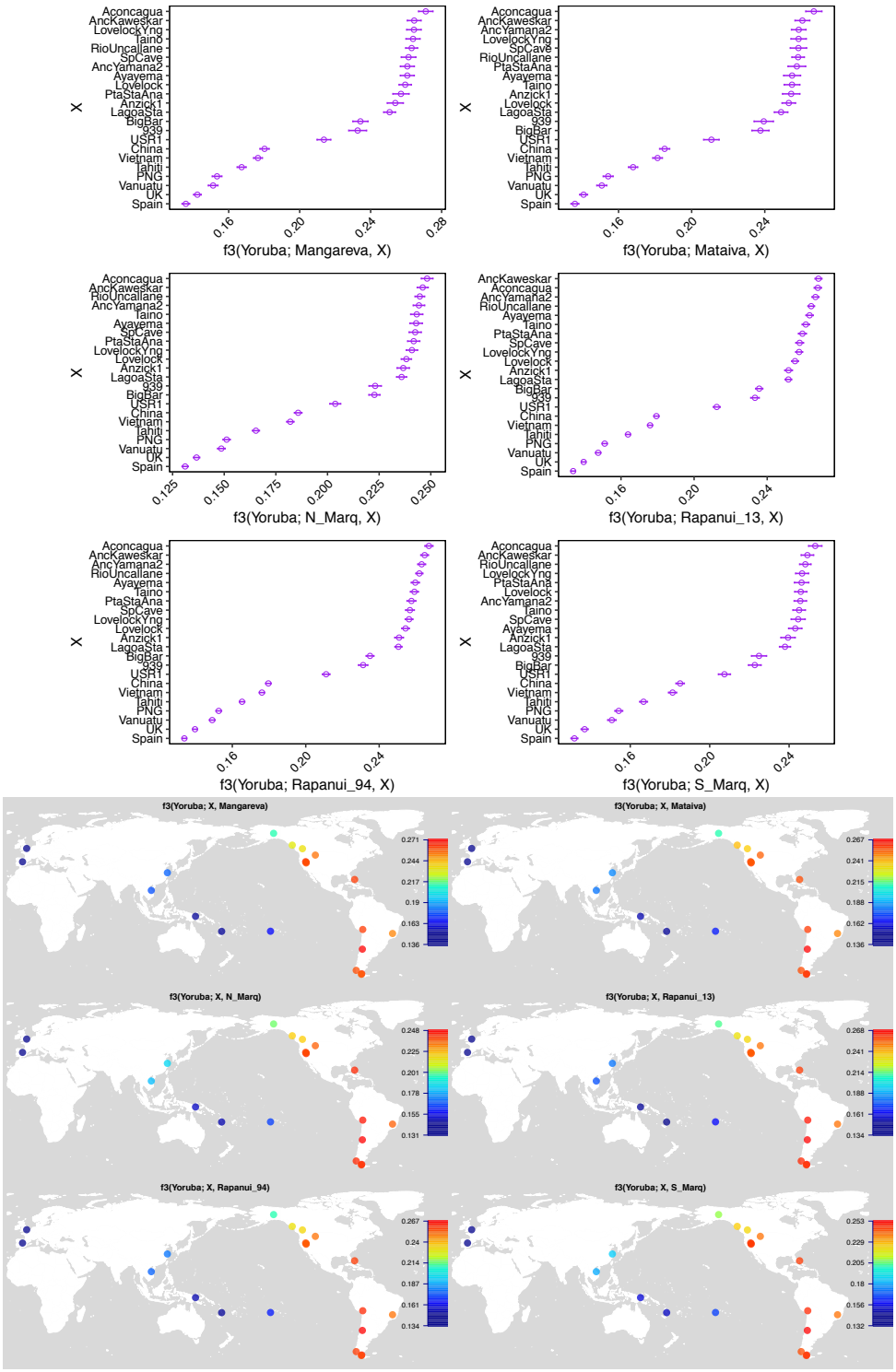
Supplementary Figure 20. Procrustes transformation applied to align two separate ASPCAs of the Native American component in Polynesians and references.

The Procrustes algorithm was used to align two separate Native American ancestry-specific PCAs created using the previously published[70] algorithm: one of samples genotyped on Illumina MEGA and Affymetrix Axiom-LAT 1 arrays (91,835 SNPs intersection) (Supplementary Figure 19) and one of samples genotyped on Illumina MEGA, Affymetrix Axiom-LAT 1, and Illumina 610-Quad arrays (28,653 SNPs three-way intersection). The Native American ancestry of reference individuals is plotted with colored points. The Native American ancestry of Pacific island individuals having greater than 2% Native American ancestry (according to local ancestry) is plotted with colored text. The centroid of the positions of the individuals from each island are plotted with stars matching the color of that island's text labels. With the exception of the Rapanui with high Native American ancestry (>40%), who cluster with the references from Chile, the Native American ancestry in Pacific island individuals is seen to cluster with northern South American and Central American references. For the low Native American ancestry individuals from the Pacific islands, the noise (scatter) due to the low number of SNPs available in the Illumina MEGA – Affymetrix Axiom LAT-1 intersection from which their coordinates derive is evident. The sample sizes for each group are given in parentheses after the group's name.



Supplementary Figure 21. Comparison of previous ancestry specific PCA (ASPCA) to our new algorithm.

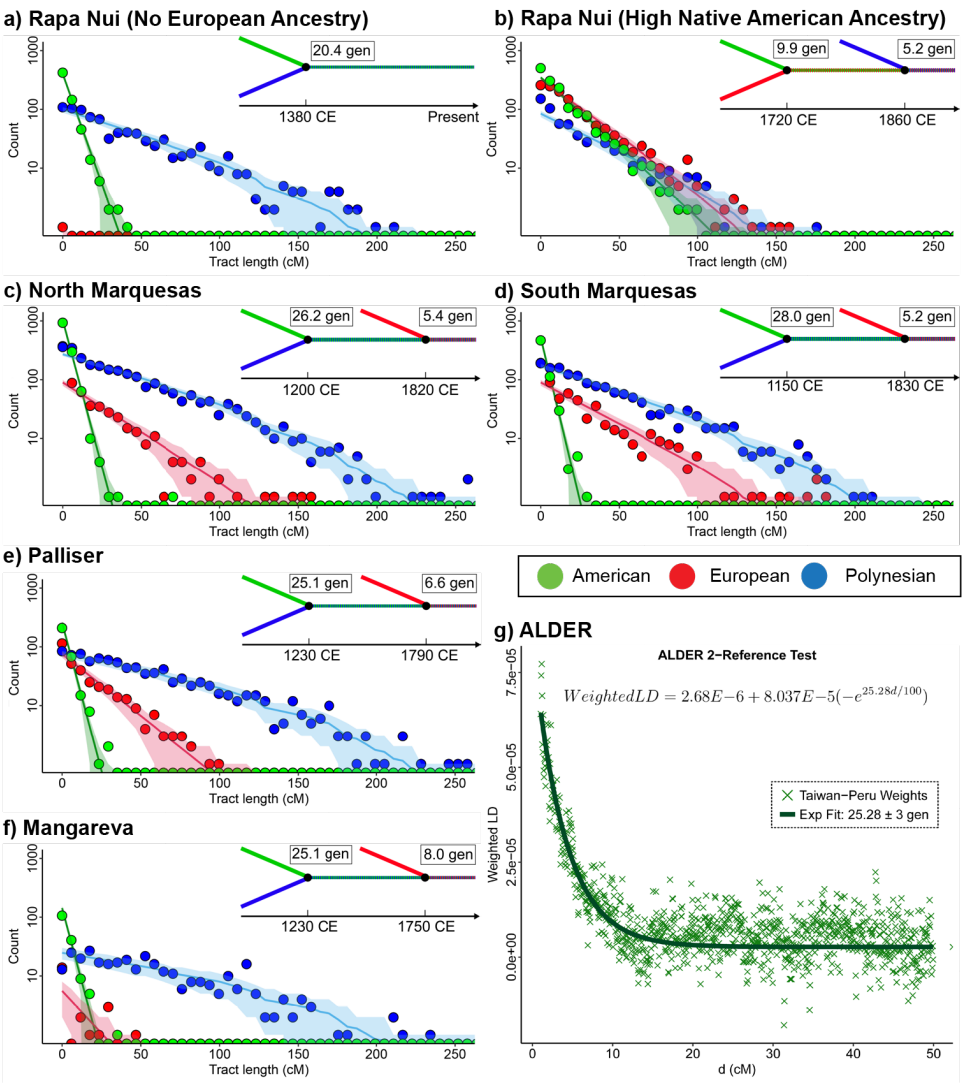
Comparison of the localization of the Native American ancestry in our dataset using the previous (left) ancestry specific PCA (ASPCA) algorithm[70] (see Supplementary Figure 19) versus our new (right) ancestry specific PCA algorithm (see Supplementary Figure 14). Southern Native American groups are depicted with yellow bars, Central Native American groups are depicted with green error bars, and the Native American ancestry in eastern Polynesian islands is depicted with light blue bars. The midline is the median and the upper and lower limits of the box are the third and first quartile respectively. The whiskers extend to 1.5 times the interquartile range from the top (or bottom) of the box to the furthest sample within that distance. Outliers beyond that distance are represented as individual points. (The full frequency distributions for these groups are shown in Figure 5b.) Native American ancestry in the Pacific islands (light blue) is seen to fall closest to the indigenous references from Colombia, the Zenu (right). This includes the "Rapa Nui (No Eur);" that is, Rapanui with no European ancestry. The "Rapa Nui (Eur)," Rapanui with high Native American ancestry (depicted as a dark blue bar), however, fall closest to the southern Native American groups (Pehuenche, Mapuche, and Huilliche) of Chile. This is consistent with the recent history of admixed Chilean individuals immigrating to Rapa Nui and contributing combined European and indigenous Chilean ancestry to inhabitants (the Rapanui). The population sample sizes for each group are given in Supplementary Data Table 1.



Supplementary Figure 22. Outgroup F3-statistics for Native American ancestry in Polynesians and the pre-contact Americas

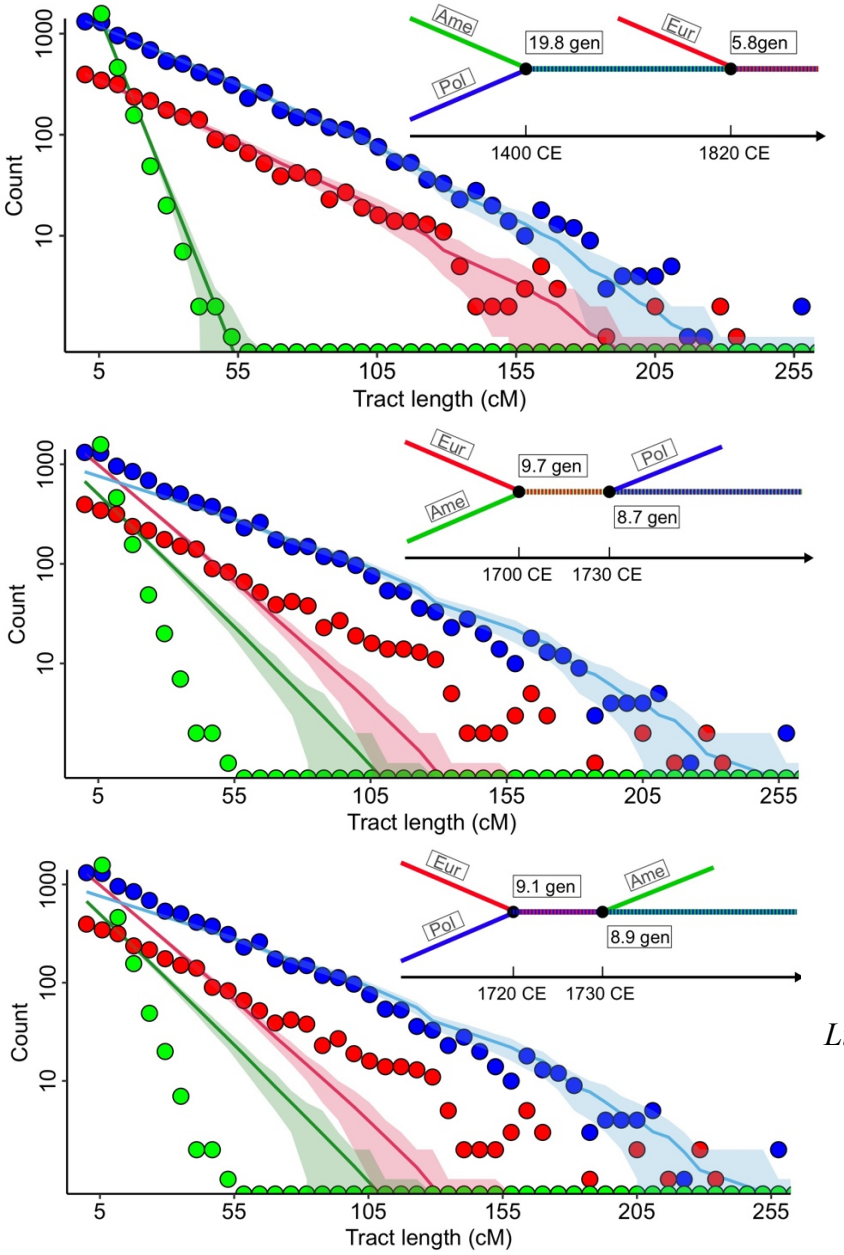
We computed outgroup F3-statistics (points) of the form *F3*(*Yoruba*; *Polynesian Target*, *X*), where we only considered sites contained in Native American ancestry tracts for each Polynesian Target population (sample sizes given in Supplementary Data Table 1), and where *X* varies across a set of reference populations (sample sizes given in Supplementary Data Table 1 and 3)

and pre-contact ancient Native American individuals (each of sample size one, see Supplementary Data Table 4). Due to the low SNP overlap of each ancient sample with the sparse Native American tracts in the Polynesians, the sensitivity of this analysis is much lower than our previous approaches for localizing the Native American ancestry found in Polynesia. Plotted standard error intervals (estimated through a weighted block-jackknife procedure over 5-Mb blocks) overlap for nearly all of the ancient samples spanning the Pacific coast of the Americas (top). In addition, there is a gap of available ancient samples in precisely our region of interest; none originate from the Pacific nations north of Peru and south of Mesoamerica. Nevertheless, we observe a general signal of genetic affinity between the Native American ancestry in Polynesians and the northernmost of the South American samples used (Aconcagua). The sharing is much weaker with samples from northern North America and Asia. The sole exception, the 2013 samples from Rapa Nui, show stronger affinity with the AncKaweskar (a Chilean population that clusters with the modern indigenous groups of central Chile in our ADMIXTURE analysis, Fig. 1). As discussed earlier, this is because European ancestry is much higher in the recent (2013) Rapanui samples, and European ancestry on Rapa Nui is correlated with the central Chilean indigenous component (Supplementary Figure 12, Supplementary Data Tables 7-8). This is likely because both are arriving together on the island, which is now part of Chile, through admixture of islanders with modern Chilean immigrants.



Supplementary Figure 23. Dating the Native American and European admixture events in Polynesia

(a-f) Tracts analyses showing the aggregate tract length counts for individuals from each island (plotted points), maximum likelihood best fit model (lines), and one standard deviation confidence intervals assuming Gaussian noise (shaded). The models corresponding to the best fit admixture chronologies are plotted as line-histories with each line colour representing each ancestry as shown in the key. Sample sizes for each group are listed in Supplementary Data Table 1. (g) Weighted LD decay curve and best fit according to the ALDER method using unadmixed indigenous Aymara as Native American reference and indigenous Taiwanese as Austronesian references (number of samples shown in Supplementary Data Table 1). All samples from islands (a-f) that were lacking European ancestry were used, namely 6 Rapanui, 4 Mangareva, 2 Palliser, and 1 North Marquesas. The discretized weighted LD was fit to an exponential curve plus an affine term using least-squares, and the standard error was found by a jack-knife across the chromosomes.



Best Model

Likelihood: **-297.86**

Likelihood: -1243.14

Likelihood: -1241.47

Supplementary Figure 24. Tract distributions for Rapanui with no southern Native American ancestry

Tracts analyses showing the aggregate tract length counts (points) of genomic regions of European (red), central Native American (green), and Polynesian (blue) ancestry in the 64 Rapanui individuals having no southern Native American ancestry. The maximum likelihood model (lines), and one standard deviation confidence intervals assuming Gaussian noise (shaded), are shown for the best fitting model for each potential sequence of admixture events with the corresponding admixture chronology plotted as a line-history inset (each line color represents an ancestry). The best fit model (at top, Native American – Polynesian admixture in 1400 CE, followed by European introgression in 1820 CE) has a higher likelihood and much better set of fit lines than the alternative admixture sequences.